

Máster en Estadística e Investigación Operativa

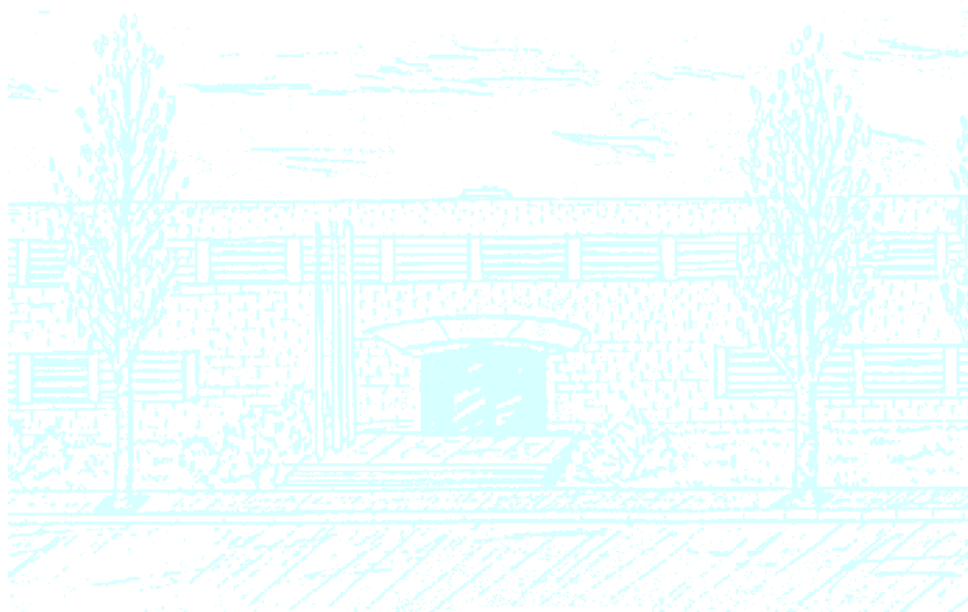
Título: Análisis longitudinal de los factores asociados a la pérdida de densidad mineral ósea en pacientes infectados con VIH

Autor: Marta Figueras Balsells

Director: Nuria Pérez Álvarez / Guadalupe Gómez Melis

Departamento: Estadística e Investigación Operativa

Convocatoria: Octubre 2009



**Facultat de Matemàtiques
i Estadística**

UNIVERSITAT POLITÈCNICA DE CATALUNYA

TRABAJO DE FIN DE MÁSTER

Análisis longitudinal de los factores asociados a la pérdida de densidad mineral ósea en pacientes infectados con VIH

por

MARTA FIGUERAS BALSELLS

MEMORIA

**Para acceder a la titulación del
Master de Estadística e Investigación Operativa por la
UNIVERSIDAD POLITÉCNICA DE CATALUÑA**

Directoras del proyecto Final de Master

Nuria Pérez Álvarez

Guadalupe Gómez Melis

Barcelona, 12 de Octubre del 2009

Índice

1	Introducción.....	1
1.1	Motivación.....	1
1.2	Conceptos básicos del VIH.....	4
2	Objetivo.....	9
3	Material y Métodos.....	11
3.1	Descripción del estudio.....	11
3.2	Descripción de las variables	11
3.2.1	Variable dependiente.....	12
3.2.2	Covariables o variables independientes	12
4	Metodología estadística.....	15
4.1	Análisis descriptivo de la población	15
4.2	Regresión lineal	15
4.2.1	Modelo lineal.....	15
4.3	Regresión logística	17
4.3.1	Modelos de respuesta binaria.....	17
4.3.2	Modelos de respuesta politómica ordinal	20
4.4	Modelos mixtos	23
4.4.1	Modelo lineal mixto	24
4.4.2	Modelos lineales generalizados mixtos	27
4.4.3	Modelos mixtos para respuesta politómica ordinal.....	28
4.5	Ecuaciones de Estimación Generalizadas. Modelos GEE	29
5	Resultados.....	31
5.1	Análisis descriptivo	31
5.2	Regresión lineal	35
5.3	Regresión logística	36
5.4	Modelo lineal mixto	37
5.5	Modelo de Ecuaciones de Estimación Generalizadas.....	38
6	Discusión y conclusiones.....	41
7	Referencias.....	43
8	Anexo	45
8.1	Covariables Sociodemográficas. Descriptiva de todos los pacientes.....	45
8.2	Covariables Bioquímicas. Descriptiva de todos los pacientes.....	46
8.3	Covariables Dietéticas. Descriptiva de todos los pacientes.....	47
8.4	Comorbilidades. Descriptiva de todos los pacientes	47
8.5	Covariables Farmacológicas. Descriptiva de todos los pacientes	48
8.6	Correlaciones bivariadas	50
8.7	Covariables Sociodemográficas. Descriptiva de los pacientes con más de una DEXA	52
8.8	Covariables Bioquímicas. Descriptiva de los pacientes con más de una DEXA	52
8.9	Comorbilidades. Descriptiva de los pacientes con más de una DEXA.....	54
8.10	Covariables Farmacológicas. Descriptiva de los pacientes con más de una DEXA	54
8.11	Modelo lineal mixto.....	56
9	Apéndice.....	59
9.1	Encuesta de factores de riesgo asociados a la osteopenia y a la osteoporosis	59

1 Introducción

1.1 Motivación

La infección por el Virus de Inmunodeficiencia Humana (VIH) es, en las últimas décadas, una de las mayores preocupaciones sociosanitarias de España; se ha situado entre las primeras causas de muerte en la población joven en los países industrializados y se ha reconocido como un problema de salud pública por su impacto humano, social, económico y mediático.

El Síndrome de Inmunodeficiencia Adquirida (SIDA) es la fase más avanzada de la enfermedad. Por consenso, la entrada en la fase del SIDA tras la infección, se diagnostica cuando el recuento de células CD4+ es menor de 200 células/mm, y/o a través de la detección de una serie de enfermedades oportunistas que se consideran indicativas de que el individuo tiene el síndrome de inmunodeficiencia adquirida y que han sido recogidas para orientar el diagnóstico. Entre estas enfermedades, la candidiasis esofágica, la neumonía causada por *pneumocystis carinii*, la toxoplasmosis cerebral, la tuberculosis o el sarcoma de kaposi son las más representativas.

Al inicio de la epidemia del SIDA, los indicadores utilizados para hacer un seguimiento del paciente y decidir sobre el tratamiento más adecuado eran la propia situación clínica del paciente y el recuento de células CD4+. En los últimos años, una nueva medida, la carga viral (número de partículas de virus que circulan en el organismo) se ha añadido a este indicador y ha adquirido el primer lugar en la investigación clínica del VIH/SIDA. Se ha demostrado que el nivel de carga viral permite pronosticar la evolución del SIDA y ayuda a predecir la mortalidad de la enfermedad, ya que las personas con una carga viral alta tienen más posibilidades de progresar rápidamente a SIDA que las personas con niveles bajos de carga viral.

En España, los primeros casos de SIDA se detectaron en los años ochenta y por desgracia, los datos de los últimos 20 años muestran que España se hallaba, junto con Portugal, entre los primeros países de Europa en términos de infectados por VIH y enfermos de SIDA, a pesar de que al inicio de la epidemia España se encontraba entre los países con menor número de afectados. Asimismo, España es de los países en que más infecciones se han dado por compartir jeringuillas entre personas con

problemas de drogadicción, siendo éstas el 60% de nuevos casos de SIDA notificados en 1993. En los últimos años se ha establecido el número de contagios en el entorno de los 5.000 ó 6.000 nuevos casos al año, y se han conseguido avances importantes en el control de la enfermedad, al tiempo que los programas de prevención y toma de conciencia han conseguido un efecto considerable.

El Tratamiento Antirretroviral de Gran Actividad, TARGA (Highly Active AntiRetroviral Treatment, HAART en inglés), ha supuesto una auténtica revolución en el curso de la enfermedad y constituye una de las mayores aportaciones humanas y sociales de la investigación farmacéutica. Los medicamentos antirretrovirales tienen el objetivo de reducir la carga viral del paciente e impedir la replicación del virus y aumentar el recuento de CD4+ para conseguir, en la medida de lo posible, una restauración del sistema inmunológico del paciente.

Los tratamientos antirretrovirales (ARV) se utilizaron primero en monoterapia (un único medicamento), más tarde se empezó a aplicar la combinación de dos fármacos y, en la actualidad, la combinación de tres fármacos antirretrovirales se ha convertido en el estándar del tratamiento del VIH/SIDA y ha supuesto un revulsivo en el abordaje terapéutico de la enfermedad por su eficacia, sin embargo, al tratarse de un tratamiento que debe mantenerse de forma continuada e indefinida, puede asociarse a efectos secundarios a largo plazo, como las alteraciones del metabolismo óseo.

El hecho de que la enfermedad se haya convertido en crónica, y que la esperanza de vida de los pacientes sea mayor, hace que la prevalencia de enfermedades geriátricas aumente, sumándose así a las posibles toxicidades generadas por el tratamiento antirretroviral, en concreto las alteraciones del metabolismo óseo (la prevalencia estimada de osteopenia y osteoporosis en pacientes infectados por el VIH es mayor a la esperada).

La osteoporosis es una enfermedad en la cual disminuye la cantidad de mineral en el hueso. La parte de hueso trabecular pierde fuerza y la zona cortical se reduce por un defecto en la absorción del calcio, lo que los vuelve quebradizos y susceptibles de fracturas y de microfracturas.

La osteopenia significa baja densidad ósea y no es una enfermedad, sino un marcador para el riesgo de fracturas.

Algunos factores de riesgo para sufrir osteopenia y osteoporosis descritos en la literatura son:

- Tiempo prolongado de la infección por el VIH.
- Tiempo prolongado en tratamiento antirretroviral.
- Tomar inhibidores de la proteasa contra la infección por el VIH.
- Bajo recuento de linfocitos CD4.
- Avanzada edad.
- Bajo peso corporal.

El trabajo que se presenta a continuación surge de la colaboración entre la Universidad Politécnica de Cataluña (UPC) y la Unidad de VIH – Fundació de Lluita contra la SIDA. La Fundació de Lluita contra la SIDA es una entidad sin ánimo de lucro fundada en el 1992 por el Dr. Bonaventura Clotet a partir de la unidad VIH del Hospital Universitari Germans Trias i Pujol (Can Ruti), que tiene como objetivos la asistencia, la investigación y la docencia en el campo de la infección del VIH.

Este proyecto tiene como objetivo determinar, mediante la aplicación de técnicas estadísticas, las características de los pacientes que indican un mayor riesgo de presentar alteraciones de la mineralización ósea en pacientes infectados por el VIH. Esta memoria consta de diferentes capítulos, el contenido de los cuales se detalla a continuación.

En el segundo capítulo se describen los objetivos del estudio. En el tercero es donde se detalla la población diana, los cuestionarios utilizados, el tipo de muestras biológicas recogidas y la definición de las variables de trabajo.

Los métodos estadísticos utilizados para llevar a cabo los análisis planteados, se encuentran en el capítulo cuatro. En este capítulo hay un primer apartado donde se hace un análisis descriptivo de la muestra. En el segundo y tercer apartado se utiliza la regresión lineal para describir la relación existente entre las covariables y la variable respuesta en un corte transversal que pretende reflejar la situación actual. Y finalmente en el cuarto apartado nuevamente se utiliza la regresión para describir la relación existente entre las covariables y la variable respuesta; que en este caso se hace de forma longitudinal con la finalidad de describir e identificar los factores asociados a los diferentes patrones de evolución en el estado de la densidad mineral ósea.

En el quinto capítulo se muestran los resultados de los diferentes análisis realizados mediante los métodos explicados en el apartado anterior. En el sexto capítulo se detallan las conclusiones tanto clínicas como estadísticas. Finalmente, se lista la bibliografía utilizada para la realización de este estudio.

1.2 Conceptos básicos del VIH

El VIH pertenece a la clase de virus llamados retrovirus. Los retrovirus son unos virus cuyo genoma está construido por ácido ribonucleico (ARN) que debe copiarse provisionalmente a ácido desoxirribonucleico (ADN) para poder multiplicarse e integrarse en la célula que infecta.

Como todos los virus, el VIH sólo puede replicarse dentro de la célula, apropiándose así de toda la maquinaria de ésta para reproducirse. Una vez dentro de la célula la conversión de ARN en ADN se lleva a cabo mediante acciones enzimáticas de transcriptasa inversa.

El VIH pertenece a un subgrupo de retrovirus conocido como lentivirus. El curso de infecciones con estos virus se caracteriza por un largo intervalo entre la infección inicial y el principio de serios síntomas.

El VIH tiene un diámetro de unos 100 nanómetros y es aproximadamente esférico. Su parte exterior es la cubierta, una membrana que originalmente pertenecía a la célula de donde el virus emergió. En la cubierta se encuentra una proteína del virus, la gp41 o glicoproteína transmembrana. Conectada a la gp41 está la gp120, la cual puede unirse al receptor CD4 localizado en la superficie de los linfocitos T para penetrar en ellos. El núcleo tiene la cápside, compuesta por la proteína p24. En su interior está el ARN.

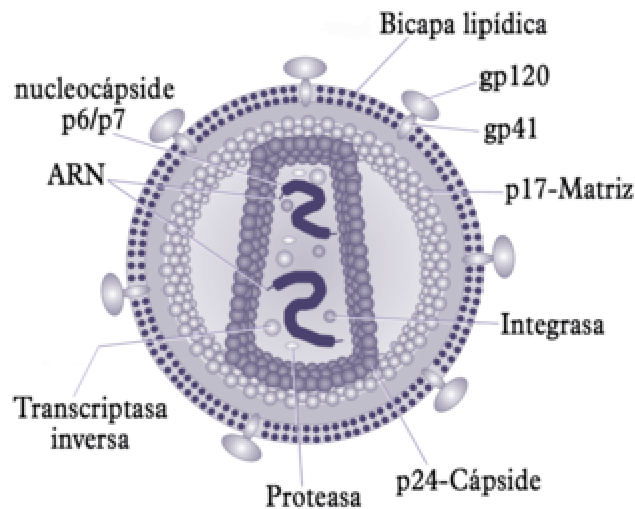


Figura 1. Estructura del virión del VIH

Las células que el VIH invade son esencialmente los linfocitos T CD4+, pero también en menor medida los monocitos/macrófagos, las células dendríticas, las células de Langerhans y las células microglía del cerebro. La replicación viral tienen pues lugar en tejidos diversos (de ganglios linfáticos, intestino, cerebro, timo,...). Los órganos linfoides, sobre todo los ganglios linfáticos, constituyen la principal sede de su replicación. El virus está presente en numerosos líquidos del organismo, en particular la sangre y las secreciones genitales.

La replicación del virus se desarrolla en las siguientes etapas:

- La fijación representa la primera etapa en la invasión de una célula. Se basa en el reconocimiento mutuo y acoplamiento de proteínas de la envoltura del virión, las gp120 y gp41, y los receptores de la célula blanco, los CD4. Este reconocimiento no es posible sin la ayuda de correceptores propios de las células susceptibles de ser invadidas; en el caso de los macrófagos son los CCR5 y en el caso de los LT4, los CXCR4, que interactúan con la proteína superficial. Macrófagos y LT4 tienen en común su principal receptor: el receptor CD4. Este reconocimiento es condición obligada para que el virus llegue a penetrar en la célula y continuar con el proceso de infección.
- La penetración es el segundo paso: una vez reconocido el virión por los receptores de superficie, se vacía dentro de la célula fusionándose la envoltura lipídica del virión con la membrana plasmática de la célula. Protegidos por la cápside y las nucleocápsides, los dos ARN mensajeros que forman el genoma viral y sus proteínas asociadas se encuentran ahora en el citoplasma.

- En el citoplasma de la célula, la transcriptasa inversa del VIH convierte el ARN viral en ADN.
- El paso siguiente es la integración del genoma vírico en el genoma de la célula huésped. Para ello penetra en el núcleo y se inserta en el ADN celular con ayuda de una integrasa, que procede del virión infectante.
- La transcripción del ADN vírico por los mecanismos normales de la célula es el quinto paso. El resultado de la transcripción es un ARNm (ARN mensajero). El ARNm obtenido es complejo, constituido por una sucesión de intrones (partes no informativas) y exones (partes informativas). Este debe ser procesado por cortes y reempalmes antes de que la información que contiene pueda servir para fabricar las proteínas correspondientes. Una vez procesado, el ARNm puede salir del núcleo a través de los poros nucleares.
- Una vez en el citoplasma el ARNm proporciona la información para la traducción, es decir, la síntesis de proteínas, que es realizada a través del aparato molecular correspondiente, del que forman la parte fundamental los ribosomas. El resultado de la traducción no consiste inmediatamente en proteínas funcionales, sino en poliproteínas que aún deben ser cortadas en fragmentos.
- Por acción de proteasas específicas del VIH, las poliproteínas producto de la traducción son procesadas, cortándolas, para formar las proteínas constitutivas del virus.
- Las proteínas víricas fabricadas se ensamblan, junto con ARN provirales, para formar los componentes internos de la estructura del virión, los que constituyen la cápside y su contenido.
- El último paso es la gemación, cuando los nucleoides víricos se aproximan a la membrana plasmática y se hacen envolver en una verruga que termina por desprenderse, formando un nuevo virión o partícula infectante. En cada célula infectada se ensamblan varios miles de nuevos viriones, aunque muchos son incompletos y no pueden infectar.

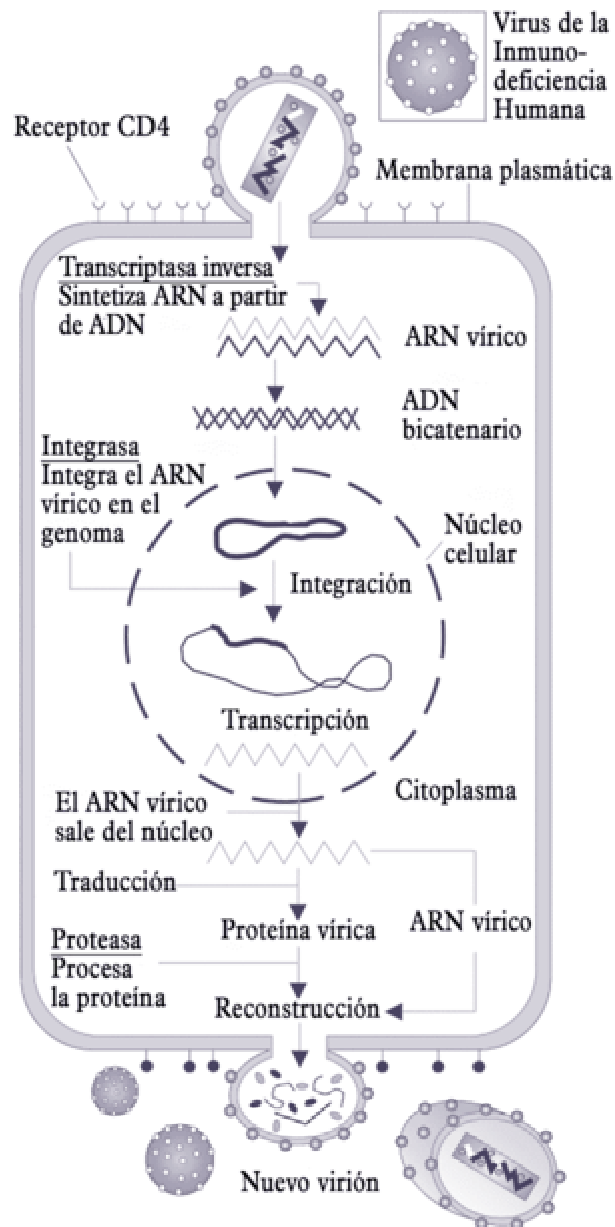


Figura 2. Ciclo de replicación del VIH

En diciembre de 2006, de acuerdo con la Organización Mundial de la Salud (OMS), había 39,5 millones de personas con VIH en el mundo, de las cuales 24,7 millones vivían en África Subsahariana.

Los fármacos antirretrovirales son medicamentos para el tratamiento de la infección por el retrovirus VIH, causante del SIDA. Diferentes antirretrovirales actúan en varias etapas del ciclo vital del VIH.

Las combinaciones de antirretrovirales actúan incrementando el número de obstáculos para la mutación viral, manteniendo bajo el número de copias. Los agentes

antirretrovirales individualmente no suprimen la infección por VIH a largo plazo, por lo cual deben usarse en combinaciones.

Los fármacos antirretrovirales existentes son:

- Inhibidores de transcriptasa inversa: Como hemos visto anteriormente la transcriptasa inversa es una enzima viral que el VIH requiere para su reproducción. La inhibición de esta enzima impide el desarrollo de ADN viral basado en su ARN.
- Inhibidores de proteasa: La proteasa es una enzima requerida por el VIH para el ensamblaje final de los viriones, por lo cual su inhibición resulta en inhibición de la replicación viral.
- Inhibidores de fusión: Impiden la entrada del virus a la célula previniendo la infección de células sanas.
- Inhibidores de la integrasa: Evitan la acción de la integrasa, selectivos únicamente al VIH. Son potentes antirretrovirales y se ha demostrado cierta eficiencia como monoterapia.
- Anticuerpos monoclonales anti-CCR5: Son anticuerpos dirigidos contra estos correceptores. Estos anticuerpos bloquean la unión de la gp120 al receptor de la célula y de esta forma el VIH no logra entrar a la célula, la célula se mantendrá intacta y continuará con sus funciones normales. Estos ofrecen una gran ventaja, ya que los actuales ARV's actúan dentro de la célula cuando el VIH ya la ha infectado. Estos anticuerpos aún no se encuentran en el mercado, pero se está estudiando su posibilidad con una aplicación clínica, como alternativa al tratamiento del VIH multirresistente, con tropismo al correceptor CCR5.

2 Objetivo

El objetivo de nuestra investigación es identificar las variables que determinan un mayor riesgo de presentar disminución de la mineralización ósea en pacientes infectados por el VIH.

Conocer los factores asociados permitirá que los clínicos tengan un mayor conocimiento del problema. Dicha información también permitirá poner en funcionamiento medidas de detección y mejores tratamientos para aquellos pacientes que tengan unas características que les hagan más susceptibles a la pérdida de densidad mineral ósea.

Los objetivos planteados han sido:

1. Describir las características de los pacientes incluidos en el estudio.
2. Describir el estado de salud de los pacientes incluidos en cuanto a la densidad mineral ósea según la categorización en normal, osteopenia y osteoporosis.
3. Establecer las relaciones entre las características demográficas y biológicas de los pacientes y la presencia de osteopenia y osteoporosis.
4. Describir la evolución en el tiempo de la densidad mineral ósea en aquellos pacientes en los que se disponga de más de una determinación.

3 Material y Métodos

3.1 *Descripción del estudio*

Este es un estudio observacional, en que los sujetos son pacientes seropositivos mayores de dieciocho años que acudieron a consulta del Hospital de Día de VIH y que disponían de al menos una DEXA (Dual-Energy X-Ray Absorptiometry).

La DEXA es una prueba radiológica que mide la densidad ósea (la cantidad de mineral óseo contenido en un cierto volumen de hueso) al pasar rayos X con dos grados diferentes de energía a través del hueso. El escáner DEXA toma mediciones repetidas para un mismo paciente en diferentes puntos del cuerpo, principalmente en la columna vertebral, cadera y muñeca. Los puntos y medidas que se recogen, y se interpretan clínicamente, han cambiado con el tiempo, debido a los nuevos hallazgos en la investigación en el área del VIH y las enfermedades óseas. Actualmente, los equipos de densitometría nos proporcionan los resultados de densidad mineral ósea en términos absolutos (g/cm^2) y relativos (T-score). El T-score es el número de desvíos estándar en que el valor absoluto medido en un determinado individuo se aleja del valor promedio de una población de individuos jóvenes, del mismo sexo, sanos, sin antecedentes de fracturas por fragilidad.

La recogida de los datos se realizó mediante la revisión de la historia clínica de los pacientes, mediante la encuesta de factores de riesgo asociados a la osteopenia y a la osteoporosis (ver apéndice) y mediante consultas en la base de datos del hospital. Estos datos fueron introducidos por un técnico en una base de datos creada con el software Access.

3.2 *Descripción de las variables*

La variable que tenemos como objetivo caracterizar es el estado óseo usando datos demográficos, antecedentes familiares, historia antirretroviral, tratamientos concomitantes, comorbilidades, estado nutricional, así como la evolución inmunológica y virológica recogida previamente o en el momento de la DEXA.

3.2.1 Variable dependiente

La DEXA es nuestra variable dependiente. Como ya se ha comentado anteriormente, el escáner DEXA toma mediciones repetidas para un mismo paciente en diferentes puntos del cuerpo, principalmente en la columna vertebral, cadera y muñeca. Estas medidas que se recogen se han resumido en una única variable categórica, según la definición de la OMS, de la siguiente manera:

- Osteoporosis: Si el mínimo de estas medidas tiene un T-score ≤ -2.5 .
- Osteopenia: Si el mínimo de estas medidas tiene un T-score entre -2.5 y -1.0.
- Normal: Si el mínimo de estas medidas tienen un T-score ≥ -1.0 .

3.2.2 Covariables o variables independientes

Las covariables elegidas en este estudio y que pueden influenciar la variable dependiente las podemos agrupar en: sociodemográficas, bioquímicas, dietéticas, comorbilidades y farmacológicas.

Es importante mencionar que el valor que toman las covariables varía dependiendo de si estamos hablando del análisis transversal o del análisis longitudinal. En el transversal se analiza el resultado de la última DEXA y para ello necesitamos conocer el valor que toman estas covariables en dicho momento o previamente, en cambio en el análisis longitudinal necesitamos conocer el valor que toman estas covariables en el momento de la primera DEXA o anterior, es decir, se seleccionan valores de las covariables anteriores en el tiempo a la medición de la respuesta.

Covariables Sociodemográficas

Las variables sociodemográficas incluyen: edad, sexo, altura, peso, índice de masa corporal (IMC), fumador, número de cigarrillos/día, práctica de ejercicio físico, su frecuencia (horas semanales) y menopausia.

Respecto a las variables fumador, número de cigarrillos/día, práctica de ejercicio físico y su frecuencia sólo se pudo obtener su valor en el momento de la última DEXA, por lo tanto no se pudieron incluir en el análisis longitudinal.

Covariables Bioquímicas

Las variables bioquímicas incluyen: CD4 nadir (número de CD4 más bajo que una persona haya tenido durante un periodo determinado o bien a lo largo de su historial médico), CD4, CD8, carga viral, creatinina, filtrado glomerular y porcentaje de tiempo con carga viral indetectable.

Para obtener el valor de las variables CD4, CD8, carga viral, creatinina y filtrado glomerular se seleccionó la determinación más próxima a la DEXA de entre el conjunto de todos los valores recogidos para un mismo paciente. Paralelamente la carga viral se clasificó en: Indetectable (≤ 50 copias/ml) o Detectable (> 50 copias/ml).

Covariables Dietéticas

Las variables dietéticas incluyen únicamente la ingesta de calcio. El valor de esta variable solamente se pudo obtener en el momento de la última DEXA, por lo que no se incluyó en el análisis longitudinal.

Comorbilidades

Estas incluyen: tiempo desde la infección por VIH, naive, hepatitis y si sufrieron alguna patología concomitante.

Respecto a la patología concomitante, se consideró que un paciente la padecía si en el momento de la DEXA sufría alguna enfermedad endocrinológica, artritis o cáncer hematológico. En el caso de la hepatitis, se eligieron todas las determinaciones previas a la DEXA del conjunto de todos los valores recuperados para un mismo paciente; si en alguno de ellos el valor era positivo, se consideraba que el paciente tenía hepatitis.

Covariables Farmacológicas

Las variables farmacológicas incluyen: tiempo en tratamiento antirretroviral, tiempo en tratamiento con inhibidores de la proteasa (IP), tiempo en tratamiento con tenofovir (un inhibidor de la transcriptasa inversa), porcentaje de tiempo con tratamiento antirretroviral con respecto al tiempo con infección por VIH, si recibían tratamiento basal, tratamiento hormonal sustitutivo, tratamiento concomitante, calcio y/o alendronato.

En cuanto al tratamiento concomitante, se consideró que un paciente lo tomaba si en el momento de la DEXA estaba siendo tratado con corticoides, antiepilépticos, quimioterapia o anticoagulantes.

Finalmente, el tratamiento basal hace referencia a si estaban recibiendo algún tipo de tratamiento antirretroviral en el momento de la DEXA; en caso afirmativo también se cuantificó cuantos de ellos estaban tratados con IP, nucleosidos o tenofovir.

4 Metodología estadística

En este apartado se describe la metodología estadística utilizada para dar respuesta a los objetivos planteados. Su aplicación ha sido posible gracias a los conocimientos adquiridos en algunas de las asignaturas del Master de estadística e investigación operativa, especialmente en Modelos lineales Generalizados y Análisis de datos longitudinales.

4.1 *Análisis descriptivo de la población*

Para conseguir un conocimiento general de las características de la población en estudio, se describirán los datos demográficos y otras características basales mediante un análisis descriptivo.

De igual modo se ha realizado un análisis descriptivo de las DEXAs en global y según datos demográficos y características basales influyentes.

En general los datos se presentan mediante estadísticos de resumen de la siguiente manera:

- Para las variables continuas: n , media, desviación estándar (DE), valor mínimo, valor máximo, mediana y percentil del 25% y del 75%.
- Para las variables categóricas: frecuencia y porcentaje de pacientes por categoría.

4.2 *Regresión lineal*

La regresión lineal es el método utilizado para estimar la relación existente entre diversas variables independientes y una variable dependiente, que en nuestro caso será continua.

4.2.1 **Modelo lineal**

Si disponemos de una colección de p variables independientes, las cuales denotaremos por la matriz $\mathbf{X}' = (X_1, X_2, \dots, X_p)$ y asumimos que las componentes de la variable dependiente $\mathbf{Y}' = (Y_1, Y_2, \dots, Y_n)$ tienen distribuciones normales, independientes, con $E[\mathbf{Y}] = \boldsymbol{\mu}$ y varianza constante σ^2 , entonces el modelo lineal viene dado por la ecuación:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i \quad i = 1, 2, \dots, n.$$

La estimación de los parámetros del modelo lineal se puede hacer mediante el método de los mínimos cuadrados o mediante el método de máxima verosimilitud, pero hay que tener en cuenta que en ausencia de normalidad, los estimadores por mínimos cuadrados no son eficientes, es decir, tienen una varianza superior a la varianza de los estimadores por máxima verosimilitud.

Una vez estimado el modelo, el interés se centra por una parte en estudiar la significación de las variables incluidas y, por otra parte en medir la bondad de ajuste del mismo.

La significación de las variables incluidas puede realizarse a partir de un estadístico distribuido según la ley de Fisher, relacionado con la variación de la suma de cuadrados residuales $SCR = (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$ y los grados de libertad del modelo, el cual recibe el nombre de contraste de la varianza incremental.

Supongamos que se quiere evaluar la significación de una variable, entonces se construyen dos modelos, el modelo 1, que no contiene la variable evaluada y el modelo 2 que contiene todas las variables que aparecen en el modelo 1 más la covariante objeto de evaluación (en esta situación se dice que el modelo 1 está anidado al 2).

Entonces se comparan las sumas de cuadrados residuales de los dos modelos:

$$F = \frac{(SCR_1 - SCR_2)/q}{SCR_2/(n-p)}.$$

Bajo la hipótesis nula, F tiene una distribución F de Fisher con grados de libertad iguales a q, n-p; siendo q la diferencia en número de parámetros entre los dos modelos, n el número de observaciones y p el número de parámetros del modelo 2.

En el caso que nos ocupa utilizaremos la estrategia stepwise para la construcción del modelo, ya que es una herramienta efectiva en el análisis de los datos⁴. El algoritmo que utiliza esta estrategia es una combinación de la estrategia forward selection y la backward elimination. La forward selection parte del modelo nulo y se van añadiendo,

una por una, aquellas covariables que satisfagan el criterio de entrada. La backward elimination parte del modelo completo y se van eliminando, una por una, aquellas covariables que satisfagan el criterio de salida.

Una medida que nos cuantifica la bondad de ajuste del modelo a los datos es el estadístico R^2 , conocido con el nombre de coeficiente de determinación.

Finalmente es necesario realizar una evaluación de los residuos del modelo final para validar las hipótesis de la regresión lineal y por tanto, garantizar las propiedades estadísticas de los estimadores del modelo asumido.

Los residuos se definen como la diferencia entre los valores observados y los valores ajustados por el modelo, $\mathbf{e} = \mathbf{Y} - \hat{\mathbf{Y}}$, y su evaluación se realiza mediante herramientas gráficas de la estadística descriptiva:

- Histograma de los residuos: se persigue ver que estén centrados en el cero y que su distribución sea aproximadamente normal.
- Normal Probability Plot de los residuos estandarizados o estudentizados.
- Diagrama bivalente de los residuos (en ordenadas) frente a los valores ajustados. La visualización de una banda horizontal de residuos centrada en $Y=0$ indica satisfacción de las hipótesis.
- Diagrama bivalente de los residuos (en ordenadas) frente a cada uno de los regresores. La visualización de una banda horizontal de residuos indica satisfacción de las hipótesis.

4.3 Regresión logística

La regresión logística es el método utilizado para estimar la relación existente entre una lista de covariables y una variable respuesta categórica.

4.3.1 Modelos de respuesta binaria

Si disponemos de una colección de p variables independientes, las cuales denotaremos por el vector $\mathbf{x}' = (x_1, x_2, \dots, x_p)$ y asumimos que la variable dependiente puede tomar los valores codificados por 0 y 1, entonces el logit del modelo de regresión logística dicotómica viene dado por la ecuación:

$$\text{logit}(\pi(\mathbf{x})) = \ln\left(\frac{\pi(\mathbf{x})}{1-\pi(\mathbf{x})}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p, \quad (\text{I})$$

Siendo $\pi(\mathbf{x}) = P(y=1|\mathbf{x})$, es decir, la probabilidad de que el resultado sea igual a 1 condicionado al vector, \mathbf{x} , de p covariables.

De la misma manera se puede escribir:

$$\pi(\mathbf{x}) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}, \quad (\text{II})$$

Lo cual significa que x_i es un factor de riesgo para y si $\beta_i > 0$ y x_i es un factor protector para y si $\beta_i < 0$. En el caso que $\beta_i = 0$, y es independiente de x_i .

Si alguna de las variables independientes, x_q por ejemplo, es discreta con las categorías $1, \dots, s$, $s > 2$, entonces variables ficticias tal como se muestran a continuación son introducidas en el modelo:

$$x_{q1} = \begin{cases} 1 & x_q = 2 \\ 0 & \text{en caso contrario} \end{cases}, \dots, x_{qs-1} = \begin{cases} 1 & x_q = s \\ 0 & \text{en caso contrario} \end{cases}$$

Siendo en este caso $x_q = 1$ el nivel de referencia.

La estimación de los parámetros del modelo logístico se hace mediante el método de máxima verosimilitud. Este método asigna valores para los parámetros desconocidos que maximizan la probabilidad de obtener el conjunto de datos observados.

Una vez estimado el modelo, el interés se centra por una parte en estudiar la significación de las variables incluidas y, por otra parte en medir la bondad de ajuste del mismo. Una medida que nos cuantifica la bondad de ajuste del modelo a los datos es la devianza y el R^2 generalizado.

Supongamos que se quiere evaluar la significación de una variable, entonces se construye el modelo 1 anidado al modelo 2 que contiene todas las variables que aparecen en el modelo 1 más la covariante objeto de evaluación.

Entonces se comparan las verosimilitudes de los dos modelos mediante la devianza:

$$D = -2\ln(\text{verosimilitud modelo 1} / \text{verosimilitud modelo 2}).$$

Bajo la hipótesis nula, D tiene una distribución Chi-cuadrado con grados de libertad iguales a la diferencia en número de parámetros entre los dos modelos, que en el caso que se acaba de describir es 1. La interpretación de la devianza es la misma que la suma de cuadrados residuales en el caso de la regresión lineal. Cuando se tenga que seleccionar un modelo entre los posibles candidatos, se seleccionará aquel que tenga una devianza más pequeña.

Igual que en el apartado anterior, la estrategia que utilizaremos para la construcción del modelo será también la estrategia stepwise, pero en este caso utilizaremos el logaritmo de la verosimilitud en vez de la suma de cuadrados residuales para la comparación de modelos y la distribución Chi-cuadrado en vez de la F de Fisher para la significación estadística.

Finalmente es necesario realizar una evaluación de los residuos del modelo final. Del mismo modo que en la regresión lineal, las medidas de las distancias entre el valor observado y el esperado son funciones de un residuo que se define como la diferencia entre el valor observado y el ajustado. En regresión logística hay distintas medidas posibles para evaluar la diferencia entre el valor observado y el ajustado. Utilizaremos el término 'patrón de covariables' para agrupar aquellos individuos que poseen el mismo valor en las covariables. Para poner de relieve el hecho que los valores ajustados en la regresión logística se calculan para cada patrón de covariables y teniendo en cuenta (II), el valor ajustado es igual a:

$$m_j \hat{\pi}_j = m_j \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)},$$

Donde m_j es el número de individuos con patrón j y $j=1,2,\dots,J$.

Dos de estas posibles medidas son el residuo de Pearson y el residuo proveniente de la devianza. Para un patrón particular de covariables el residuo de Pearson se define como:

$$r(y_j^*, \hat{\pi}_j) = \frac{(y_j^* - m_j \hat{\pi}_j)}{\sqrt{m_j \hat{\pi}_j (1 - \hat{\pi}_j)}},$$

Siendo y_j^* la suma de la variable respuesta para los individuos del patrón j -ésimo.

El estadístico de resumen basado en estos residuos, el estadístico Chi-cuadrado de Pearson, se define como:

$$\chi^2 = \sum_{j=1}^J r(y_j^*, \hat{\pi}_j)^2 .$$

Análogamente, los residuos devianza se definen como:

$$d(y_j^*, \hat{\pi}_j) = \pm \left\{ 2 \left[y_j^* \ln \left(\frac{y_j^*}{m_j \hat{\pi}_j} \right) + (m_j - y_j^*) \ln \left(\frac{m_j - y_j^*}{m_j (1 - \hat{\pi}_j)} \right) \right] \right\} ,$$

Y el estadístico de resumen basado en estos residuos, el estadístico devianza, se define como:

$$D = \sum_{j=1}^J d(y_j^*, \hat{\pi}_j)^2 .$$

Bajo la hipótesis de que el ajuste es correcto, ambos estadísticos se distribuyen asintóticamente como una Chi-cuadrado con $J-(p+1)$ grados de libertad; donde p corresponde al número de variables independientes y J al número de perfiles o patrones distintos observados. Estos residuos pueden ser representados gráficamente versus los valores ajustados para evaluar posibles valores anómalos.

4.3.2 Modelos de respuesta politómica ordinal

En el apartado anterior nos hemos centrado en el caso del modelo de regresión logística cuando la variable respuesta es binaria, pero el modelo puede ser fácilmente modificado para el caso donde la variable respuesta tiene k niveles, $k > 2$, que pueden ser nominales u ordinales.

En este apartado nos centraremos en los modelos para escalas ordinales puesto que es el tipo de variable que debemos estudiar.

En muchas aplicaciones, la definición de las categorías de la variable de respuesta es arbitraria o subjetiva, sin embargo, resulta fundamental que las conclusiones del estudio sean válidas y no dependan ni del número ni de la definición de las categorías de la respuesta, de manera que si se forma una nueva escala en base a crear una

nueva categoría combinando dos categorías adyacentes de la antigua escala, la esencia de las conclusiones debe permanecer.

Las anteriores consideraciones llevan a pensar directamente en modelos basados en las probabilidades acumuladas de las respuestas $\gamma_j = P(y \leq j)$, no en las probabilidades mismas $\pi_j = P(y=j)$.

Los dos conjuntos de probabilidades son equivalentes, pero los modelos basados en probabilidades acumuladas parece que deban tener mejores propiedades para las escalas ordinales. En particular, los modelos que emplean la transformación logística sobre las probabilidades acumuladas han demostrado funcionar bien en la práctica⁶.

Los modelos más simples dentro de la tipología anterior son los que definen rectas paralelas en la escala de los log-odds acumulados,

$$\ln\left(\frac{\gamma_j(x)}{1-\gamma_j(x)}\right) = \alpha_j - \beta^T x \quad j=1,2, \dots, k-1.$$

Donde $\gamma_j(x) = P(y \leq j | x)$, es decir, la probabilidad de que el resultado sea menor o igual a j condicionado al vector, x , de p covariables.

Este modelo es conocido como modelo de los odds proporcionales, ya que el odds-ratio del suceso $y \leq j$ para $x=x_1$ y $x=x_2$ es,

$$\frac{\gamma_j(x_1)/(1-\gamma_j(x_1))}{\gamma_j(x_2)/(1-\gamma_j(x_2))} = e^{-\beta^T(x_1-x_2)} \quad j=1,2, \dots, k-1$$

el cual es independiente de la elección de la categoría j .

También es posible formular modelos más complejos de rectas no paralelas, donde el resultado será la estimación de $k-1$ rectas de regresión no paralelas,

$$\ln\left(\frac{\gamma_j(x)}{1-\gamma_j(x)}\right) = \alpha_j - \beta_j^T x \quad j=1,2, \dots, k-1.$$

Seleccionamos el modelo que define rectas paralelas al demostrar, mediante los resultados obtenidos a partir de los dos modelos logísticos binarios (variable dependiente categorizada como osteoporosis y osteopenia/normal y como osteoporosis/osteopenia y normal), un buen ajuste de nuestros datos.

El signo negativo en los parámetros β es un convenio para garantizar que los valores grandes de $\beta^T x$ faciliten probabilidades elevadas a las categorías de orden superior. Los parámetros α y β son las incógnitas a estimar, con la restricción de que los términos independientes de las rectas no sean crecientes, es decir, $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_{k-1}$.

La estimación de los parámetros para este modelo se hace mediante el método de máxima verosimilitud.

Una vez más la estrategia que utilizaremos para la construcción del modelo será la de stepwise, descrita en el apartado anterior.

Igual que en el apartado de los modelos de respuesta binaria, es necesario medir la bondad de ajuste del mismo, pero cabe destacar que en este caso este estudio es notablemente más difícil, al tener un resultado de más de dos niveles, pues cuando la variable respuesta tiene tres categorías tenemos dos probabilidades logísticas estimadas. Lesaffre (1986) y Lesaffre and Albert (1989) han presentado extensiones de test de bondad de ajuste para modelos de respuesta politómica, sin embargo, estos métodos no son fácilmente calculables mediante el software disponible. Por lo tanto, mientras que los softwares no añadan estos métodos a sus paquetes, lo más óptimo es calcular la bondad de ajuste de los dos modelos de regresión logística por separado y luego tomar una decisión sobre la bondad de ajuste del modelo final considerando cada uno de los resultados.

Por último, propondremos algunos ejemplos relacionados con la interpretación de los parámetros del modelo logístico y politómico.

Supongamos que queremos estudiar la variable DEXA mediante los tres modelos siguientes:

- Modelo 1 - categorías variable respuesta Osteoporosis y Osteopenia/Normal:

$$\ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \alpha_1 + \beta \cdot \text{Sexo}, \text{ donde } \pi(x) = \text{Prob. de Osteoporosis}$$
- Modelo 2 - categorías variable respuesta Osteoporosis/Osteopenia y Normal:

$$\ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \alpha_2 + \beta \cdot \text{Sexo}, \text{ donde } \pi(x) = \text{Prob. de Osteoporosis u Osteopenia}$$
- Modelo 3 - categorías de la variable respuesta Osteoporosis, Osteopenia y Normal:

$$\ln\left(\frac{\gamma_1(x)}{1-\gamma_1(x)}\right) = \alpha_1 + \beta \cdot \text{Sexo}, \text{ donde } \gamma_1(x) = \text{Prob. de Osteoporosis}$$

$$\ln\left(\frac{y_2(x)}{1-y_2(x)}\right) = \alpha_2 + \beta \cdot \text{Sexo}, \text{ donde } y_2(x) = \text{Prob. de Osteoporosis u Osteopenia}$$

Y supongamos que los resultados de la estimación son:

Modelo	α_1	α_2	β (Categoría de referencia: Mujer)
1	-1.2769		0.3093
2		0.5222	0.7710
3	-1.4894	0.6412	0.5792

Tabla 1. Estimación de los parámetros.

La estimación de α_i indica el odds ratio en la categoría de referencia. La interpretación de la covariable sexo sería la siguiente:

- Modelo 1: El odds ratio de osteoporosis en los hombres es 0.3093 veces el odds ratio que presentan las mujeres.
- Modelo 2: El odds ratio de osteoporosis/osteopenia en los hombres es 0.7710 veces el odds ratio que presentan las mujeres.
- Modelo 3: El odds ratio de disminución de la densidad mineral ósea en los hombres es 0.5792 veces el odds ratio que presentan las mujeres.

Las estimaciones de α_i son muy parecidas para los dos modelos de respuesta binaria y el politómico que asume rectas paralelas.

4.4 Modelos mixtos

Los modelos mixtos se caracterizan por permitir modelar datos cuyas observaciones no son independientes, aceptando así la existencia de correlaciones entre los datos de unos determinados subgrupos de la muestra.

Existen varios tipos de modelos mixtos. En los modelos de efectos fijos las covariables y factores considerados para explicar la variabilidad de la variable respuesta se presuponen constantes y por lo tanto los parámetros del modelo toman valores constantes. En los modelos de efectos aleatorios se asume que algunos parámetros son realizaciones de valores que provienen de una muestra independiente, la cual sigue una distribución normal. Finalmente, en los modelos mixtos con medidas repetidas, las medidas son tomadas repetidamente sobre el mismo individuo y además

de los parámetros descritos por el modelo de efectos fijos o por el modelo de efectos aleatorios se puede añadir un parámetro indicando el efecto del tiempo.

Algunas de las ventajas de los modelos mixtos son:

- Si escogemos una buena estructura de covarianza, podemos obtener unas estimaciones más apropiadas de los efectos fijos y de los errores estándar.
- Cuando los datos tienen una estructura jerárquica, los resultados pueden ser más precisos.
- Podemos elegir varianzas distintas para diferentes grupos de sujetos.
- Podemos evitar los problemas causados por los datos faltantes al estimar los parámetros.

4.4.1 Modelo lineal mixto

Como ya hemos visto en el apartado 4.2.1, los modelos de efectos fijos pueden ser expresados como:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} + \mathbf{e},$$

$$\mathbf{V} = \text{var}(\mathbf{y}) = \sigma^2 \mathbf{I},$$

Donde

$\mathbf{y} = (y_1, y_2, y_3, \dots, y_n)'$ = valores observados,

$\boldsymbol{\alpha} = (\mu, \alpha_1, \alpha_2, \dots, \alpha_p)'$ = efectos fijos,

$\mathbf{e} = (e_1, e_2, e_3, \dots, e_n)'$ = residuos,

σ^2 = varianza residual,

\mathbf{I} = matriz identidad $n \times n$,

\mathbf{X} = matriz de diseño $n \times p$,

\mathbf{V} = matriz de varianzas/covarianzas de las observaciones.

El modelo lineal mixto extiende el modelo de efectos fijos incluyendo efectos aleatorios, coeficientes aleatorios y/o patrones de covarianzas entre las observaciones. Estos pueden ser expresados como:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} + \mathbf{Z}\boldsymbol{\beta} + \mathbf{e},$$

Donde \mathbf{y} , \mathbf{X} , $\boldsymbol{\alpha}$ y \mathbf{e} son definidos de la misma manera que en el modelo de efectos fijos y

$\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \dots, \beta_q)'$ = efectos / coeficientes aleatorios,

Z = segunda matriz de diseño n x q.

La covarianza de **y**, $\text{var}(\mathbf{y}) = \mathbf{V}$, puede ser escrita como:

$$\mathbf{V} = \text{var}(\mathbf{X}\boldsymbol{\alpha} + \mathbf{Z}\boldsymbol{\beta} + \mathbf{e}).$$

Ahora, al asumir que los efectos aleatorios y los residuos son incorrelacionados tenemos que,

$$\mathbf{V} = \text{var}(\mathbf{X}\boldsymbol{\alpha}) + \text{var}(\mathbf{Z}\boldsymbol{\beta}) + \text{var}(\mathbf{e}).$$

Como $\text{var}(\mathbf{X}\boldsymbol{\alpha}) = \mathbf{0}$, pues $\boldsymbol{\alpha}$ describe los efectos fijos, y **Z** es una matriz de constantes resulta que:

$$\mathbf{V} = \mathbf{Z} \text{var}(\boldsymbol{\beta}) \mathbf{Z}' + \text{var}(\mathbf{e}).$$

Finalmente, si denotamos **G** a la $\text{var}(\boldsymbol{\beta})$ y **R** a la $\text{var}(\mathbf{e})$, pues como los efectos aleatorios y los residuos siguen una distribución normal podemos escribir $\boldsymbol{\beta} \sim N(\mathbf{0}, \mathbf{G})$ y $\mathbf{e} \sim N(\mathbf{0}, \mathbf{R})$, se llega a:

$$\mathbf{V} = \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R}.$$

La dimensión de **G** es q x q, donde q es el número total de parámetros aleatorios. En los modelos de efectos aleatorios, G es siempre diagonal, es decir se asume que los efectos aleatorios son incorrelacionados.

Por otro lado, al asumir también que en los modelos de efectos aleatorios los residuos son incorrelacionados, tenemos que $\mathbf{R} = \sigma^2 \mathbf{I}$.

La estimación de los parámetros del modelo lineal mixto se hace mediante el método de máxima verosimilitud. En los modelos donde las observaciones son independientes, la función de verosimilitud es el producto de las funciones de densidad de cada observación. Sin embargo, en un modelo mixto, como las observaciones no son independientes, la función de verosimilitud necesita ser basada en una función de densidad multivariante que puede ser definida usando la distribución normal multivariante, siendo entonces:

$$L = \frac{\exp\left[-\frac{1}{2}(\mathbf{Y} - \mathbf{X}\boldsymbol{\alpha})' \mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\alpha})\right]}{(2\pi)^{(1/2)n} |\mathbf{V}|^{1/2}}.$$

En la práctica, habitualmente, se utiliza el logaritmo de la función de verosimilitud en vez de la función de verosimilitud, siendo este:

$$\log(L) = K - \frac{1}{2} \left[\log|\mathbf{V}| + (\mathbf{Y} - \mathbf{X}\boldsymbol{\alpha})' \mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\alpha}) \right],$$

Donde

$$K = -\frac{1}{2}n \log(2\pi) \text{ (una constante que puede ser ignorada en el proceso de maximización),}$$

n = número de observaciones.

Existen varios métodos para ajustar el modelo lineal mixto, todos ellos basados en la maximización la función de verosimilitud. Estos métodos son:

- Máxima verosimilitud (ML).
- Máxima verosimilitud residual (REML).
- Iterativo de mínimos cuadrados generalizados (IGLS).

El método de máxima verosimilitud (ML) consiste en, primero maximizar el logaritmo de la función de verosimilitud con respecto a los parámetros de la varianza, considerando los efectos fijos α como constantes y después, una vez obtenidas las estimaciones de los parámetros de la varianza, obtener la estimación de los efectos fijos considerando los parámetros de la varianza como constantes. Este método tiene el efecto de producir estimaciones de los parámetros de la varianza sesgadas, siendo este sesgo mayor cuanto menores sean los grados de libertad.

En el método de máxima verosimilitud residual (REML), conocido también como método de máxima verosimilitud restringida, el parámetro α es eliminado del logaritmo de la función de verosimilitud siendo definido sólo en términos de los parámetros de la varianza de la siguiente manera:

$$\log(L(\mathbf{y}; \mathbf{y} - \mathbf{X}\hat{\alpha})) = K - \frac{1}{2} \left[\log|\mathbf{V}| - \log|\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}|^{-1} + (\mathbf{y} - \mathbf{X}\hat{\alpha})' \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X}\hat{\alpha}) \right],$$

Siendo $\hat{\alpha} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} \mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$.

Este método al tener en cuenta que α es un parámetro y no una constante, las estimaciones de los parámetros de la varianza son no sesgadas. De igual modo que en el método ML, α es estimado considerando los parámetros de la varianza como fijos y encontrando el valor del α que maximiza la función de verosimilitud REML.

El método iterativo de mínimos cuadrados generalizados (IGLS) obtiene la estimación de los parámetros de los efectos fijos, α , minimizando el producto $(\mathbf{y} - \mathbf{X}\alpha)' \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X}\alpha)$ y la estimación de la varianza de los parámetros mediante la

solución de las ecuaciones $(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\alpha}})(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\alpha}})' = \mathbf{V}$. Este método puede ser usado iterativamente obteniendo el mismo resultado que obtendríamos usando el ML.

4.4.2 Modelos lineales generalizados mixtos

Los modelos lineales generalizados mixtos permiten ampliar los modelos lineales generalizados con efectos fijos incluyendo efectos aleatorios y patrones de covarianzas entre las observaciones.

Estos modelos se pueden definir como:

$$\mathbf{y} = \boldsymbol{\mu} + \mathbf{e},$$

Donde $\boldsymbol{\mu}$ es el vector de medias esperadas, el cual está funcionalmente relacionado con los parámetros del modelo a través de la denominada función de link g :

$$g(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\alpha} + \mathbf{Z}\boldsymbol{\beta}.$$

\mathbf{X} y \mathbf{Z} son las matrices de diseño de los efectos fijos y aleatorios, respectivamente, y $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$ son los vectores de parámetros de los efectos fijos y aleatorios, respectivamente. Si asumimos que $\boldsymbol{\beta}$ sigue una distribución normal, entonces $\boldsymbol{\beta} \sim N(\mathbf{0}, \mathbf{G})$.

La varianza de la variable \mathbf{y} viene dada por la matriz:

$$\text{var}(\mathbf{y}) = \mathbf{V} = \text{var}(\boldsymbol{\mu}) + \mathbf{R},$$

Donde \mathbf{R} es la matriz de varianzas residuales, $\text{var}(\mathbf{e})$. Sin embargo, en este caso la matriz \mathbf{V} no puede ser especificada tan fácilmente como en el caso de datos normales, ya que $\boldsymbol{\mu}$ no es una función lineal de $\boldsymbol{\beta}$. Una primera aproximación usada por algunos métodos de ajuste es:

$$\mathbf{V} \approx \mathbf{BZGZ}'\mathbf{B} + \mathbf{R},$$

Siendo $\mathbf{B} = \text{diag}\{(\mathbf{g}')^{-1}(\boldsymbol{\mu})\}$, es decir, una matriz diagonal de varianzas de la distribución subyacente.

Una de las formas más comunes de ajustar los modelos lineales generalizados mixtos es mediante la maximización de la función de verosimilitud. Sin embargo, una de las dificultades es que la verdadera función de verosimilitud sólo puede ser definida para los efectos aleatorios; no existe una verdadera función de verosimilitud para los modelos de respuesta no-normal con patrones de covarianzas entre las observaciones. De todas formas, podemos superar este problema definiendo una función alternativa conocida como función de cuasi-verosimilitud, la cual posee varias propiedades similares a la verdadera función de verosimilitud. Cabe destacar que se

puede dar una forma general de función de cuasi-verosimilitud apropiada para todo tipo de modelos mixtos.

Actualmente existen varios métodos para maximizar la función de cuasi-verosimilitud, de los cuales no se puede recomendar ninguno como el mejor para ajustar todos los modelos lineales generalizados mixtos³.

4.4.3 Modelos mixtos para respuesta politómica ordinal

Los modelos mixtos para respuesta politómica ordinal son una ampliación de los modelos de respuesta politómica ordinal, vistos en el apartado 4.3.2, utilizados para poder incluir efectos aleatorios y patrones de covarianzas entre las observaciones.

El modelo mixto ordinal se puede definir como:

$$\begin{aligned} \mathbf{y} &= \boldsymbol{\mu} + \mathbf{e}, \\ \log(\boldsymbol{\mu}^{[c]} / (1 - \boldsymbol{\mu}^{[c]})) &= \mathbf{X}\boldsymbol{\alpha} + \mathbf{Z}\boldsymbol{\beta}, \\ \boldsymbol{\beta} &\sim N(\mathbf{0}, \mathbf{G}), \\ \text{var}(\mathbf{e}) &= \mathbf{R}, \end{aligned}$$

Donde

$$\boldsymbol{\mu} = (\mu_{11}, \mu_{12}, \mu_{13}, \mu_{21}, \mu_{22}, \mu_{23}, \dots, \mu_{n1}, \mu_{n2}, \mu_{n3})'$$

μ_{ij} = probabilidad de que la observación i esté en la categoría j ,

$$\boldsymbol{\mu}^{[c]} = (\mu_{11}^{[c]}, \mu_{12}^{[c]}, \mu_{13}^{[c]}, \mu_{21}^{[c]}, \mu_{22}^{[c]}, \mu_{23}^{[c]}, \dots, \mu_{n1}^{[c]}, \mu_{n2}^{[c]}, \mu_{n3}^{[c]})',$$

$$\mu_{ij}^{[c]} = P(y_i \leq j) = \sum_{k=1}^j \mu_{ik}.$$

Cabe destacar que la matriz \mathbf{R} de varianzas residuales tienen en cuenta las correlaciones multinomiales entre los vectores binarios usados para cada observación y cualquier patrón de covarianza definido a nivel residual.

Llegados a este punto, como el modelo está en la forma de un modelo lineal generalizado mixto, se puede definir una función de cuasi-verosimilitud y usar los mismos métodos para maximizar tal función del mismo modo que hemos expuesto en el apartado anterior.

Otra alternativa para ajustar estos modelos es incorporar los efectos aleatorios al término de error en vez de en la parte lineal del modelo. Teniendo en cuenta esta especificación se puede escribir como:

$$\begin{aligned} \mathbf{y} &= \boldsymbol{\mu} + \mathbf{A}^{1/2} \mathbf{B}^{1/2} (\mathbf{Z}\boldsymbol{\beta} + \mathbf{e}), \\ g(\boldsymbol{\mu}) &= \mathbf{X}\boldsymbol{\alpha}, \\ \mathbf{Z}\boldsymbol{\beta} + \mathbf{e} &\sim N(\mathbf{0}, \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{P}) \end{aligned}$$

\mathbf{G} es una matriz que tiene la misma forma que en el modelo de efectos aleatorios y \mathbf{P} es una matriz de correlaciones conteniendo las covarianzas de los parámetros y si es igual a la identidad, \mathbf{I} , significa que no se han modelado los patrones de covarianzas.

Aunque los efectos aleatorios no se hayan incluido en la parte lineal del modelo, puesto que se asume su incorporación en el término de error, la matriz de covarianzas tiene la siguiente estructura:

$$\mathbf{V} = \mathbf{A}^{1/2} \mathbf{B}^{1/2} (\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{P}) \mathbf{B}^{1/2} \mathbf{A}^{1/2}$$

Esta alternativa no se ha utilizado en este trabajo ya que no está implementada en la mayoría de los softwares estadísticos.

4.5 Ecuaciones de Estimación Generalizadas. Modelos GEE

En el método GEE, las ecuaciones que se tienen que resolver para encontrar una estimación de los parámetros fijos son:

$$\sum_{i=1}^N \frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}} (\mathbf{A}_i(\boldsymbol{\beta})^{1/2} \mathbf{R}_i(\rho) \mathbf{A}_i(\boldsymbol{\beta})^{1/2})^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_i) = \mathbf{0},$$

Donde \mathbf{A}_i es la matriz formada por las varianzas marginales en la diagonal principal y ceros en el resto de sitios y \mathbf{R}_i es la matriz de correlaciones, conocida como "working correlation matrix".

Las matrices de correlaciones utilizadas habitualmente son las que se muestran a continuación en la Tabla 2 y que es posible incorporar un parámetro de sobredispersión, Φ , a $\mathbf{A}_i^{1/2} \mathbf{R}_i \mathbf{A}_i^{1/2}$ para poder aumentar o disminuir la varianza.

Estructura	Corr(y_{ij}, y_{ik})	Estimador
Identidad	0	-
Intercambiable	ρ	$\hat{\rho} = \frac{1}{N} \sum_{i=1}^N \frac{1}{n_i(n_i - 1)} \sum_{j \neq k} e_{ij} e_{ik}$
Autorregresiva	$\rho^{ j-k }$	$\hat{\rho} = \frac{1}{N} \sum_{i=1}^N \frac{1}{n_i - 1} \sum_{j \leq n_i - 1} e_{ij} e_{ij+1}$
Sin estructura	ρ_{jk}	$\hat{\rho}_{jk} = \frac{1}{N} \sum_{i=1}^N e_{ij} e_{ik}$

Tabla 2. Opciones más comunes para la matriz de correlaciones.

Ahora, para resolver las ecuaciones citadas anteriormente, se utiliza el siguiente procedimiento iterativo:

1. Se calculan unas estimaciones iniciales para β , $\beta^{(0)}$, usando los modelos de respuesta binaria vistos en el apartado 4.3.1.
2. Se calculan los residuos de Pearson $e_{ij} = \frac{y_{ij} - \mu_{ij}}{\sqrt{v(\mu_{ij})}}$, ρ y el parámetro de dispersión Φ

$$= \frac{1}{N} \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^{n_i} e_{ij}^2.$$

3. Se calcula $R_i(\rho)$ y $V_i = \Phi A_i(\beta)^{1/2} R_i(\rho) A_i(\beta)^{1/2}$.
4. Finalmente se estiman los parámetros mediante la ecuación

$$\beta^{(t+1)} = \beta^{(t)} - \left[\sum_{i=1}^N \left(\frac{\partial \mu_i}{\partial \beta'} \right) V_i^{-1} \left(\frac{\partial \mu_i}{\partial \beta'} \right)' \right]^{-1} \times \left[\sum_{i=1}^N \left(\frac{\partial \mu_i}{\partial \beta'} \right) V_i^{-1} (y_i - \mu_i) \right].$$

Y se iteraran los pasos 2, 3 y 4 hasta que se da la convergencia, obteniendo así una estimación de los coeficientes de las variables en el modelo.

5 Resultados

En este apartado se presentan los resultados de aplicar la metodología estadística utilizada en este proyecto y descrita en el capítulo 4. El análisis se ha realizado con el paquete estadístico SAS versión 9.0.

5.1 *Análisis descriptivo*

Un total de 671 pacientes seropositivos que se visitan en la Unidad de VIH – Fundació de Lluita contra la SIDA fueron incluidos en este estudio.

La última DEXA estudiada aquí se realizó entre los años 2000 y 2009.

En la Tabla 3 podemos ver que de los 671 pacientes incluidos, el 72% eran hombres y sólo el 1% era naïve en cuanto al hecho de no haber recibido tratamiento antirretroviral.

Variable		Total n(%)
Sexo	Mujer	188 (28.02%)
	Hombre	483 (71.98%)
Naïve	Ha tomado tratamiento	664 (98.96%)
	No ha tomado tratamiento	7 (1.04%)

Tabla 3. Distribución por sexo y naïve.

La edad mediana fue de 43 años, con un rango que iba de los 20 a los 75 años. Con respecto al índice de masa corporal (IMC) sólo se pudo conseguir el valor de 639 pacientes siendo la media de 23.37 kg/m² (ver Tabla 4).

	Edad (años)	IMC (kg/m²)
n	671	639
Media (DE)	42.86 (8.61)	23.37 (3.40)
Mediana	42.10	23.09
(Min, Max)	(20.10, 74.90)	(15.16, 43.15)
(P ₂₅ , P ₇₅)	(37.40, 47.60)	(21.05, 25.19)

Tabla 4. Descriptiva de Edad e IMC en el momento de la última DEXA.

En cuanto a la variables relacionadas con el VIH, descritas en la Tabla 5, tenemos que la media de tiempo con infección por el VIH fue de 10.94 años (DE:5.75), y con tratamiento antirretroviral fue de 7.72 años (DE:4.70). En el momento de la última DEXA, la cifra mediana de linfocitos CD4 fue de 496 células/mm³ (P₂₅:368; P₇₅:686) y la de la carga viral del VIH fue de 3.91 logaritmos (P₂₅:3.91; P₇₅:4.48). El 61% de los pacientes tuvo una carga viral indetectable (≤ 50 copias/ml).

	Tiempo con infección por el VIH (años)	Tiempo con tratamiento arv. (años)	Linfocitos CD4 (células/mm³)	Logaritmo Carga Viral
n	642	645	627	641
Media (DE)	10.94 (5.75)	7.716 (4.701)	551.05 (288.38)	5.12 (2.36)
Mediana	11.14	7.439	496.00	3.91
(Min, Max)	(0.11, 25.37)	(0.003, 33.199)	(4.41, 2025.0)	(3.00, 14.40)
(P ₂₅ , P ₇₅)	(6.80, 15.12)	(4.104, 11.034)	(368.00, 686.00)	(3.91, 4.48)

Tabla 5. Descriptiva de las variables relacionadas con el VIH.

De los 671 pacientes casi el 50% (n=323) presentaba osteopenia, un 25% (n=174) osteoporosis y el 25% (n=174) restante no mostró ninguno de estos síntomas.

De todos los pacientes reclutados sólo 391 dispusieron de más de una DEXA (ver Figura 3).

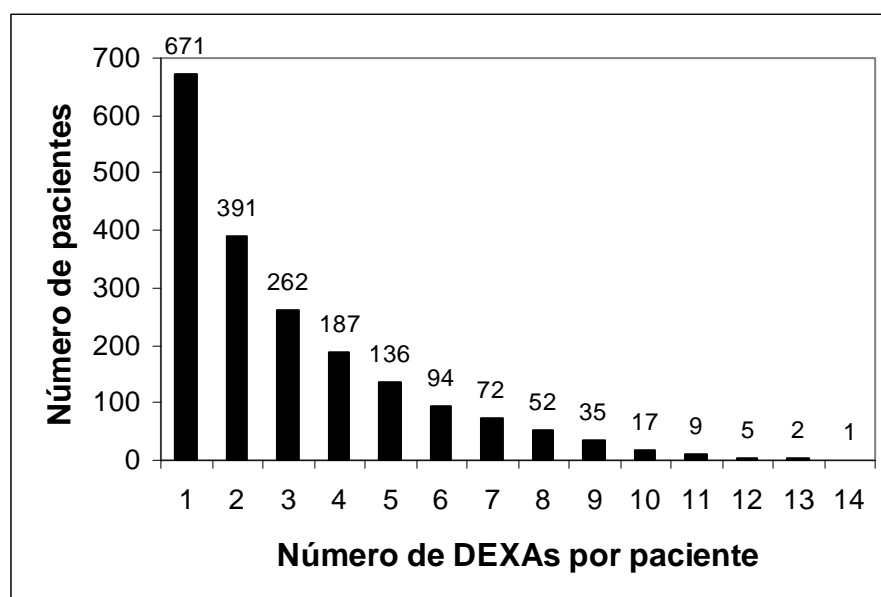


Figura 3. Número de DEXAs registradas para un mismo paciente

En la Tabla 6 podemos ver que de estos 391 pacientes, casi el 73% eran hombres y sólo dos de ellos eran naive.

Variable		Total n(%)
Sexo	Mujer	106 (27.11%)
	Hombre	285 (72.89%)
Naive	Ha tomado tratamiento	389 (99.49%)
	No ha tomado tratamiento	2 (0.51%)

Tabla 6. Distribución por sexo y naive de los pacientes con más de una DEXA.

La edad mediana fue de 43 años, con un rango que iba de los 20 a los 75 años. Con respecto al índice de masa corporal (IMC) solo se pudo conseguir el valor de 381 pacientes siendo la media de 23.16 kg/m² (ver Tabla 7).

	Edad (años)	IMC (kg/m²)
n	391	381
Media (DE)	43.56 (8.84)	23.33 (3.36)
Mediana	42.70	23.16
(Min, Max)	(20.10, 74.90)	(15.16, 36.95)
(P ₂₅ , P ₇₅)	(37.80, 48.40)	(21.03, 25.34)

Tabla 7. Descriptiva de Edad e IMC de los pacientes con más de una DEXA.

En cuanto a la variables relacionadas con el VIH, descritas en la Tabla 8, tenemos que la media de tiempo con infección por el VIH fue de 11.34 años (DE:5.75), y con tratamiento antirretroviral fue de 8.4 años (DE:4.91). En el momento de la última DEXA, la cifra mediana de linfocitos CD4 fue de 501 células/mm³ (P₂₅:373; P₇₅:699) y la de la carga viral del VIH fue de 3.91 logaritmos (P₂₅:3.91; P₇₅:4.38). El 66% de los pacientes tuvo una carga viral indetectable (≤ 50 copias/ml).

	Tiempo con infección por el VIH (años)	Tiempo con tratamiento arv. (años)	Linfocitos CD4 (células/mm³)	Log Carga Viral
n	380	386	375	385
Media (DE)	11.34 (5.75)	8.403 (4.912)	564.11 (286.51)	4.90 (2.12)
Mediana	11.78	8.451	501.00	3.91
(Min, Max)	(0.39, 25.37)	(0.504, 33.199)	(16.30, 1938.0)	(3.00, 13.22)
(P ₂₅ , P ₇₅)	(6.84, 15.79)	(4.427, 11.943)	(373.00, 699.00)	(3.91, 4.38)

Tabla 8. Descriptiva de las variables relacionadas con el VIH de los pacientes con más de una DEXA.

La media de tiempo entre una DEXA y la siguiente fue de unos 10 meses y en la Figura 4 podemos ver los perfiles de 20 pacientes seleccionados al azar.

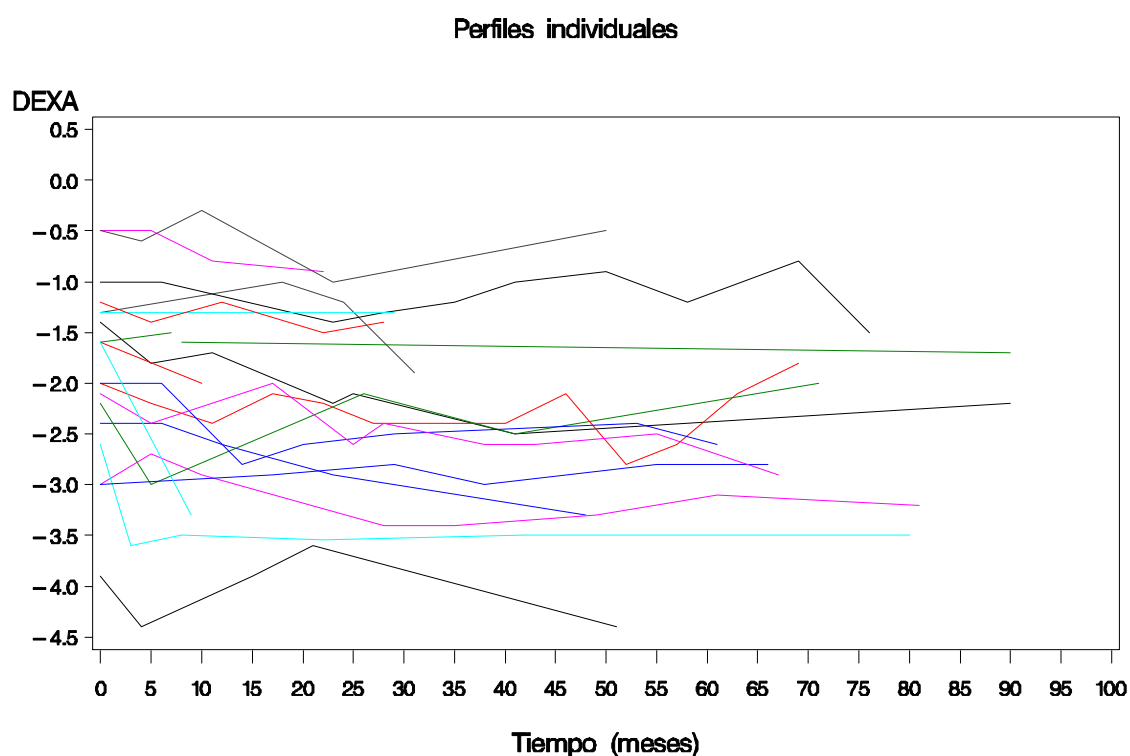


Figura 4. Perfiles individuales de 20 pacientes seleccionados al azar.

Esta figura muestra la evolución en el valor de DEXA a lo largo del tiempo. El número de DEXAs no es el mismo para todos los pacientes y tampoco se realizaron con la misma frecuencia, ni en un mismo paciente, ni en el conjunto de pacientes, lo cual nos lleva a tener datos no balanceados y medidas que no son equidistantes.

5.2 Regresión lineal

En este apartado se lleva a cabo un modelo de regresión lineal para encontrar las variables que determinan un mayor riesgo de presentar disminución de la mineralización ósea en pacientes infectados por el VIH. Los datos analizados representan un corte transversal realizado en la última DEXA de cada paciente; los datos registrados se engloban en el periodo que va de Marzo 2000 a Enero 2009.

La Tabla 9 contiene el resultado del modelo lineal final en el cual vemos que los factores de riesgo a padecer osteoporosis u osteopenia son ser hombre, de edad avanzada, con bajo IMC, con una larga exposición a los tratamientos antirretrovirales especialmente a los inhibidores de la proteasa, tomando tratamiento antirretroviral en el momento de la última DEXA y con un nivel alto de creatinina.

Parámetro		Estimación	Error estándar	P-valor
Término independiente		-0.4646	0.5453	0.3946
Edad		-0.0526	0.0062	<0.0001
IMC		0.0971	0.0155	<0.0001
Sexo (Categoría de referencia: Mujer)	Hombre	-0.3343	0.1252	0.0078
Tiempo con IP		-0.0837	0.0192	<0.0001
Tto. arv. en el momento de la última DEXA (Categoría de referencia: No)	Sí	-0.5944	0.1752	0.0007
Creatinina		-0.0101	0.0036	0.0047
Tiempo con tto. arv. (Categoría de referencia: < 25%)	>= 75%	0.7029	0.2432	0.0040
	[50 - 75)	0.6146	0.2534	0.0156
	[25 - 50)	0.5958	0.2595	0.0221

Tabla 9. Estimación de los parámetros para el modelo lineal

La interpretación de estas covariables sería la siguiente: en el caso de ser hombre la densidad mineral ósea disminuye en -0.3343 comparado con las mujeres manteniendo los otros factores constantes y el hecho de cumplir un año más, la densidad mineral ósea disminuye en -0.0526 manteniendo los otros factores constantes.

5.3 Regresión logística

La regresión logística cuantifica la asociación entre un factor de riesgo y una variable categórica. En este caso nos interesamos en que factores están asociados con la presencia de osteoporosis y osteopenia.

Como se describió en el apartado 4.3.2 esta técnica nos permite modelar la relación lineal entre el logaritmo de los odds de disminución de la densidad mineral ósea y las diferentes variables en el modelo.

En la siguiente tabla, podemos ver que las variables asociadas a padecer osteoporosis u osteopenia son muy parecidas a las del modelo lineal. Ser hombre, de edad avanzada, con bajo IMC, con una larga exposición a los tratamientos antirretrovirales, especialmente a los inhibidores de la proteasa y a tenofovir y coinfectados con hepatitis fueron identificados como factores de riesgo a padecer osteoporosis u osteopenia.

Parámetro		OR	IC 95%	P-valor
Término independiente osteoporosis		0.4450	(0.1142; 1.7346)	0.2435
Término independiente osteopenia		6.0345	(1.5423; 23.6107)	0.0098
Edad		1.0736	(1.0512; 1.0965)	<0.0001
IMC		0.8513	(0.8084; 0.8965)	<0.0001
Tiempo con IP		1.0814	(1.0098; 1.1581)	0.0251
Sexo (Categoría de referencia: Mujer)	Hombre	3.1078	(2.0953; 4.6094)	<0.0001
Hepatitis (Categoría de referencia: Sí)	No	0.6369	(0.4440; 0.9138)	0.0143
Toma IP en el momento de la última DEXA (Categoría de referencia: Sí)	No	0.4742	(0.3196; 0.7036)	0.0002
Toma TDF en el momento de la última DEXA (Categoría de referencia: Sí)	No	0.6932	(0.4947; 0.9713)	0.0332

Tabla 10. Estimación de los parámetros para el modelo logístico

En este caso, por ejemplo, tenemos que el odds ratio de disminución de la densidad mineral ósea se multiplica por 1.0736 al cumplir un año más y que el odds ratio de disminución de la densidad mineral ósea en los hombres es 3.1078 veces el odds que presentan las mujeres.

La interpretación de los dos términos independientes es la siguiente: manteniendo todos factores constantes, el odds ratio de padecer osteoporosis frente a no padecerla se multiplica por 0.4450 y el odds ratio de padecer osteoporosis u osteopenia frente a no padecer ninguno de estos dos síntomas, se multiplica por 6.0345.

5.4 *Modelo lineal mixto*

En este apartado, a fin de poder estudiar la disminución de la densidad mineral ósea a lo largo del tiempo, la muestra se ha estratificado según el año en que se realizó la primera DEXA, ya que los tratamientos antirretrovirales han evolucionado considerablemente a lo largo de los años. El hecho de estratificar por el año en que se hizo la primera DEXA también nos resultó favorable computacionalmente, bajando el tiempo de cálculo hasta la convergencia.

Es importante mencionar que en todos estos modelos se ha incluido el paciente como efecto aleatorio. El efecto tiempo permanece en el modelo aunque no sea significativo, ya que la estructura de los datos se definió así.

El modelo más relevante clínicamente fue el del año 2001, por tener un tamaño mostral más grande y al no existir tanta disparidad de pacientes con respecto al tratamiento antirretroviral.

Los factores de riesgo a padecer osteoporosis u osteopenia son: ser hombre, de edad avanzada, con bajo IMC y con un nivel alto de carga viral (ver Tabla 11).

	Estimación	Error estándar	P-valor
Efecto aleatorio	0.8431	0.1238	<0.0001
Efectos fijos			
Término independiente	-1.0626	0.7567	0.1631
Tiempo	-0.0038	0.0012	0.0012
Sexo (Categoría de referencia: Mujer) Hombre	-0.6955	0.2231	0.0019
Edad	-0.0436	0.0111	<0.0001
IMC	0.0962	0.0330	0.0037
Un logaritmo de Carga viral	-0.0769	0.0361	0.0337

Tabla 11. Estimación de los parámetros para el modelo lineal mixto. Primera DEXA año 2001.

A tiempo 0, la densidad mineral ósea de un hombre de mediana edad con 23.16 kg/m² de IMC y 3.91 log de carga viral, sería de -1.7, la cual disminuiría en -0.0769 por el hecho de aumentar en 1 log su carga viral, manteniendo el resto de factores constantes.

5.5 *Modelo de Ecuaciones de Estimación Generalizadas*

En este apartado se quiere estudiar la presencia o ausencia de padecer osteoporosis u osteopenia, teniendo en cuenta que los pacientes tienen más de una DEXA y que las observaciones de un mismo paciente pueden estar correlacionadas.

En la Tabla 12 se pueden encontrar los resultados del modelo obtenido, en el cual vemos que los factores de riesgo a padecer osteoporosis u osteopenia son: ser hombre, de edad avanzada, con bajo IMC, con una larga exposición a los tratamientos antirretrovirales, especialmente a los inhibidores de la proteasa y al tenofovir, y tomando IP en el momento de la última DEXA.

Parámetro		OR	IC 95%	P-valor
Término independiente osteoporosis		0.1720	(0.0752; 0.3934)	<0.0001
Término independiente osteopenia		1.7986	(0.7906; 4.0915)	0.1616
Tiempo		1.0105	(1.0062; 1.0147)	<0.0001
Edad		1.0693	(1.0563; 1.0824)	<0.0001
Tiempo con IP		1.1833	(1.1236; 1.2462)	<0.0001
IMC		0.8732	(0.8448; 0.9024)	<0.0001
Sexo (Categoría de referencia: Mujer)	Hombre	2.2291	(1.7619; 2.8202)	<0.0001
Tiempo con TDF		1.0872	(1.0313; 1.1463)	0.0019
Toma IP en el momento de la última DEXA (Categoría de referencia: Sí)	No	0.6091	(0.4975; 0.7457)	<0.0001

Tabla 12. Estimación de los parámetros para el modelo GEE

En este caso por ejemplo, observamos que el odds ratio de disminución de la densidad mineral ósea, es 1.0693 veces más al cumplir un año más; además el odds ratio de disminución de la densidad mineral ósea de los hombres equivale al odds de las mujeres multiplicado por 2.6781.

6 Discusión y conclusiones

Los datos recogidos han sido descritos y analizados tomando medidas en un momento concreto y tomando todas las observaciones longitudinalmente. Hemos procedido de esta manera, con el fin de determinar las variables asociadas a un mayor riesgo de presentar alteraciones de la mineralización ósea, en pacientes infectados por el VIH.

Los modelos utilizados han sido:

- Modelo lineal y modelo politómico ordinal, para el corte transversal.
- Modelo lineal mixto y modelo de ecuaciones de estimación generalizada, para el análisis longitudinal.

Los resultados obtenidos a partir de los modelos transversales nos indican que, para nuestra población, los factores de riesgo a padecer osteoporosis u osteopenia son: ser hombre, de edad avanzada, con bajo IMC, con una larga exposición a los tratamientos antirretrovirales especialmente a los inhibidores de la proteasa y a tenofovir, tomando tratamiento antirretroviral en el momento de la última DEXA, especialmente los inhibidores de la proteasa y tenofovir, con hepatitis y con un nivel alto de creatinina.

Los resultados obtenidos a partir de los modelos longitudinales son similares a los obtenidos mediante los modelos transversales siendo, en este caso, los factores de riesgo a padecer osteoporosis u osteopenia: ser hombre, de edad avanzada, con bajo IMC, con una larga exposición a los tratamientos antirretrovirales, especialmente a los inhibidores de la proteasa y al tenofovir, tomando IP en el momento de la última DEXA y con un nivel alto de carga viral.

Puede sorprender que ser hombre sea un factor de riesgo comparado con las mujeres, pero está descrito y justificado que es así en pacientes VIH+ cuando las mujeres no son menopaúsicas.

El hecho de que un nivel alto de creatinina sea un factor de riesgo es debido a que un nivel alto de creatinina significa que el paciente tiene una alteración renal y si esta alteración es del filtrado, eliminando en exceso minerales como calcio y sodio, provoca una disminución de la densidad mineral ósea.

Como los resultados obtenidos a partir de los modelos longitudinales son muy similares a los obtenidos a través de los modelos transversales, podemos decir que en este tipo de datos somos capaces de entender la evolución de la pérdida y ganancia de la densidad mineral ósea de un paciente, conociendo el valor de los factores de riesgo en el momento en que se tomó la última DEXA, sin necesidad de conocer toda su evolución. Una explicación razonable a este hecho puede ser la escasa variabilidad del parámetro a lo largo del tiempo (ver Figura 4); su último valor puede ser un buen indicador.

De esta forma podemos decir que, aunque con los modelos mixtos con medidas repetidas se obtienen resultados más precisos, en este caso parece no compensar el gran esfuerzo en la recogida de datos y la gran actividad computacional. Por este motivo recomendaría el uso de estadísticos resumen para la evolución y modelos para observaciones independientes para este tipo de datos.

7 Referencias

1. Agresti, A. (1990). *Categorical Data Analysis*, Wiley, Inc., New York.
2. Badía, X. *La aportación de los medicamentos antirretrovirales al tratamiento del VIH/SIDA*. EL valor del medicamento. Farmaindustria.
3. Brown, H., and Prescott, R. *Applied Mixed Models in Medicine*, Second Edition. John Wiley & Sons, Ltd.
4. Hosmer, D. W., and Lemeshow, S. *Applied Logistic Regression*, Second Edition. New York. John Wiley & Sons, cop. 2000.
5. JAMA, La Revista de la Asociación Médica Americana. *La osteopenia y la prevención de fracturas*. 6 de Diciembre de 2006 - Vol. 296, No. 21.
6. McCullagh, P., and Nelder, J.A. (1989). *Generalized Linear Models*, Second Edition. Chapman & Hall, London.
7. Molenberghs, G., and Verbeke, G. *Models for Discrete Longitudinal Data*. Springer.
8. Neter, J., Kutner, M. H., Nachtsheim, C. J., and Wasserman, W. *Applied Linear Statistical Models*, Chicago.Irwin, cop. 1996.
9. SAS Institute Inc., Cary, NC, USA. *SAS Help and Documentation*, Software SAS System for Windows Version 9.0.
10. Verbeke, G., and Molenberghs, G. *Linear Mixed Models for Longitudinal Data*. Springer.
11. Jacobson, DL., Spiegelman, D., Knox, TK., Wilson, IB. *Evolution and predictors of change in total bone mineral density over time in HIV-infected men and women in the nutrition for healthy living study*. J. Acquir Immune Defic Syndr. 2008 Nov 1; 49(3): 298-308.
12. Bachrach, LK., Hastie, T., Wang, MC., Narasimhan, B. Marcus, R. *Bone mineral acquisition in healthy Asian, Hispanic, Black, and Caucasian youth: a longitudinal study*. J. Clin. Endocrinol Metab. 1999 Dec; 84(12):4702-12.
13. Cazanave, C., Dupon, M., Lavignolle-Aurillac, V., Barthe, N., Lawson-Ayayi, S., Mehse, N., Mercié, P., Morlat, P., Thiébaud, R., Dabis, F.; Groupe d'Epidémiologie Clinique du SIDA en Aquitaine. *Reduced bone mineral density in HIV-infected patients: prevalence and associated factors*. AIDS. 2008 Jan 30; 22(3): 395-402.

8 Anexo

8.1 Covariables Sociodemográficas. Descriptiva de todos los pacientes

		Total
Fumador	No	373 (56.43%)
	Sí	288 (43.57%)
Número cigarrillos / día	n	281
	Media (DE)	19.47 (11.99)
	Mediana	20.00
	(Min, Max)	(1.00, 80.00)
	(P ₂₅ , P ₇₅)	(10.00, 20.00)

Tabla 13. Hábitos tóxicos de todos los pacientes en el momento de la última DEXA.

		Total
Ejercicio físico	No	499 (75.72%)
	Sí	160 (24.28%)
Horas a la semana	n	145
	Media (DE)	5.10 (2.90)
	Mediana	5.00
	(Min, Max)	(1.00, 14.00)
	(P ₂₅ , P ₇₅)	(3.00, 7.00)

Tabla 14. Ejercicio físico de todos los pacientes en el momento de la última DEXA.

		Total
Menopausia	No	169 (90.37%)
	Sí	18 (9.63%)
Tiempo con menopausia (días)	n	18
	Media (DE)	2009.39 (1411.47)
	Mediana	2295.00
	(Min, Max)	(136.00, 4651.0)
	(P ₂₅ , P ₇₅)	(548.00, 2979.0)

Tabla 15. Menopausia de todos los pacientes.

8.2 Covariables Bioquímicas. Descriptiva de todos los pacientes

		Total
Variable		
CD4 Nadir	n	650
	Media (DE)	230.54 (156.16)
	Mediana	219.00
	(Min, Max)	(1.00, 1025.0)
	(P ₂₅ , P ₇₅)	(112.00, 319.00)

Tabla 16. CD4 Nadir de todos los pacientes.

		Total
Variable		
CD4 Absoluto	n	627
	Media (DE)	551.05 (288.38)
	Mediana	496.00
	(Min, Max)	(4.41, 2025.0)
	(P ₂₅ , P ₇₅)	(368.00, 686.00)
CD4 Porcentaje	n	636
	Media (DE)	26.10 (9.12)
	Mediana	26.00
	(Min, Max)	(1.00, 58.00)
	(P ₂₅ , P ₇₅)	(20.00, 32.00)
CD8 Absoluto	n	622
	Media (DE)	1048.66 (486.87)
	Mediana	992.50
	(Min, Max)	(115.00, 3408.0)
	(P ₂₅ , P ₇₅)	(700.00, 1297.0)
CD8 Porcentaje	n	632
	Media (DE)	49.07 (11.32)
	Mediana	49.00
	(Min, Max)	(21.00, 81.00)
	(P ₂₅ , P ₇₅)	(40.00, 57.00)

Tabla 17. Linfocitos CD4 y CD8 de todos los pacientes en el momento de la última DEXA.

		Total
Variable		
Creatinina	n	627
	Media (DE)	85.24 (15.69)
	Mediana	83.50
	(Min, Max)	(48.00, 172.70)
	(P ₂₅ , P ₇₅)	(74.00, 94.20)

Tabla 18. Creatinina de todos los pacientes.

		Total
Variable		
Filtrado glomerular	<=60	1 (4.55%)
	>60	21 (95.45%)

Tabla 19. Filtrado glomerular de todos los pacientes.

		Total
Variable		
Porcentaje de tiempo con CV indetectable	< 75%	551 (86.64%)
	>= 75%	85 (13.36%)

Tabla 20. Porcentaje de tiempo con CV indetectable de todos los pacientes.

8.3 Covariables Dietéticas. Descriptiva de todos los pacientes

		Total
Variable		
Ingesta calcio	n	207
	Media (DE)	927.50 (378.93)
	Mediana	897.00
	(Min, Max)	(1.00, 2570.0)
	(P ₂₅ , P ₇₅)	(700.00, 1200.0)

Tabla 21. Ingesta de calcio de todos los pacientes.

8.4 Comorbilidades. Descriptiva de todos los pacientes

		Total
Variable		
Hepatitis	No	373 (64.20%)
	Sí	208 (35.80%)

Tabla 22. Hepatitis de todos los pacientes.

		Total
Variable		
Patología concomitante	No	625 (93.98%)
	Sí	40 (6.02%)

Tabla 23. Patología concomitante de todos los pacientes.

8.5 Covariables Farmacológicas. Descriptiva de todos los pacientes

		Total
Variable		
Tiempo con IP (años)	n	645
	Media (DE)	3.65 (2.93)
	Mediana	3.08
	(Min, Max)	(0.00, 13.23)
	(P ₂₅ , P ₇₅)	(1.34, 5.69)

Tabla 24. Tiempo con IP de todos los pacientes en el momento de la última DEA.

		Total
Variable		
Tiempo con tenofovir (años)	n	645
	Media (DE)	1.98 (2.46)
	Mediana	1.21
	(Min, Max)	(0.00, 29.14)
	(P ₂₅ , P ₇₅)	(0.00, 3.11)

Tabla 25. Tiempo con tenofovir de todos los pacientes en el momento de la última DEXA.

		Total
Variable		
Porcentaje de tiempo con tratamiento antirretroviral	< 25%	35 (5.69%)
	[25 - 50)	97 (15.77%)
	[50 - 75)	151 (24.55%)
	>= 75%	332 (53.98%)

Tabla 26. Porcentaje de tiempo con tratamiento antirretroviral de todos los pacientes.

		Total
Variable		
Tratamiento hormonal sustitutivo	No	659 (99.10%)
	Sí	6 (0.90%)

Tabla 27. Tratamiento hormonal sustitutivo de todos los pacientes.

		Total
Variable		
Tratamiento concomitante	No	583 (87.67%)
	Sí	82 (12.33%)

Tabla 28. Tratamiento concomitante (Corticoides, antiepilépticos, quimioterapia y anticoagulantes) de todos los pacientes.

		Total
Variable		
Calcio	No	644 (96.84%)
	Sí	21 (3.16%)

Tabla 29. Complemento de calcio de todos los pacientes.

		Total
Variable		
Alendronato	No	648 (97.44%)
	Sí	17 (2.56%)

Tabla 30. Alendronato de todos los pacientes.

8.6 Correlaciones bivariadas

Coeficientes de correlación Pearson Prob > r suponiendo H0: Rho=0 Número de observaciones														
	Edad	IMC	CD4 Nadir	Tiempo VIH (años)	Tiempo con tratamiento antirretroviral (años)	Tiempo con IP (años)	Tiempo con tenofovir (años)	CD4 Absoluto	CD4 Porcentaje	CD8 Absoluto	CD8 Porcentaje	Carga viral	Ingesta calcio	Creatinina
Edad	1.00000 0.0004 671	0.13897 0.0004 639	-0.19536 <.0001 650	0.29482 <.0001 642	0.37210 <.0001 645	0.26004 <.0001 645	0.23351 <.0001 645	-0.07179 0.0724 627	-0.15152 0.0001 636	0.08892 0.0266 622	0.07570 0.0572 632	-0.06710 0.0896 641	0.07759 0.2664 207	0.06569 0.1003 627
IMC	0.13897 0.0004 639	1.00000 0.0060 639	0.11028 0.0060 620	-0.08994 0.0261 612	-0.06121 0.1294 615	0.03740 0.3545 615	0.04388 0.2773 615	0.02657 0.5170 597	-0.01853 0.6488 606	0.04931 0.2305 593	0.01936 0.6352 603	-0.02194 0.5886 610	0.07460 0.2877 205	0.08745 0.0325 598
CD4 nadir	-0.19536 <.0001 650	0.11028 0.0060 620	1.00000 0.0060 650	-0.20365 <.0001 636	-0.18931 <.0001 636	-0.25545 <.0001 636	-0.06167 0.1203 636	0.51007 <.0001 627	0.46784 <.0001 636	0.08583 0.0323 622	-0.26782 <.0001 632	0.02057 0.6035 640	0.04478 0.5218 207	-0.01989 0.6191 627
Tiempo VIH (años)	0.29482 <.0001 642	-0.08994 0.0261 612	-0.20365 <.0001 636	1.00000 0.0261 642	0.70471 <.0001 629	0.43646 <.0001 629	0.31279 <.0001 629	-0.02129 0.5988 613	-0.01451 0.7180 622	0.04359 0.2833 608	0.02319 0.5651 618	-0.11539 0.0038 629	0.00820 0.9076 203	-0.01847 0.6472 616
Tiempo con tratamiento antirretroviral (años)	0.37210 <.0001 645	-0.06121 0.1294 615	-0.18931 <.0001 636	0.70471 <.0001 629	1.00000 <.0001 645	0.60107 <.0001 645	0.50082 <.0001 645	0.04304 0.2870 614	0.03038 0.4494 622	0.03946 0.3310 609	-0.06750 0.0936 618	-0.19937 <.0001 630	0.05763 0.4188 199	-0.00669 0.8685 616
Tiempo con IP (años)	0.26004 <.0001 645	0.03740 0.3545 615	-0.25545 <.0001 636	0.43646 <.0001 629	0.60107 <.0001 645	1.00000 0.3545 645	0.16189 <.0001 645	0.01197 0.7671 614	-0.08224 0.0403 622	0.11386 0.0049 609	0.03279 0.4158 618	-0.13529 0.0007 630	0.08762 0.2185 199	0.07191 0.0745 616
Tiempo con tenofovir (años)	0.23351 <.0001 645	0.04388 0.2773 615	-0.06167 0.1203 636	0.31279 <.0001 629	0.50082 <.0001 645	0.16189 <.0001 645	1.00000 <.0001 645	-0.00754 0.8521 614	0.04909 0.2215 622	-0.07372 0.0691 609	-0.11615 0.0038 618	-0.15850 <.0001 630	0.04757 0.5046 199	-0.06146 0.1276 616
CD4 Absoluto	-0.07179 0.0724 627	0.02657 0.5170 597	0.51007 <.0001 627	-0.02129 0.5988 613	0.04304 0.2870 614	0.01197 0.7671 614	-0.00754 0.8521 614	1.00000 0.5170 627	0.64403 <.0001 627	0.34056 <.0001 622	-0.41253 <.0001 623	-0.19693 <.0001 619	0.05204 0.4734 192	-0.01256 0.7574 607
CD4 Porcentaje	-0.15152 0.0001 636	-0.01853 0.6488 606	0.46784 <.0001 636	-0.01451 0.7180 622	0.03038 0.4494 622	-0.08224 0.0403 622	0.04909 0.2215 622	0.64403 <.0001 627	1.00000 0.0001 636	-0.35132 <.0001 622	-0.72828 <.0001 632	-0.26876 <.0001 628	-0.05163 0.4690 199	-0.11692 0.0037 616
CD8 Absoluto	0.08892 0.0266 622	0.04931 0.2305 593	0.08583 0.0323 622	0.04359 0.2833 608	0.03946 0.3310 609	0.11386 0.0049 609	-0.07372 0.0691 609	0.34056 <.0001 622	-0.35132 <.0001 622	1.00000 0.56647 622	0.56647 <.0001 622	0.13708 0.0007 614	0.02636 0.7196 188	0.10041 0.0137 602

Coefficientes de correlación Pearson Prob > r suponiendo H0: Rho=0 Número de observaciones														
	Edad	IMC	CD4 Nadir	Tiempo VIH (años)	Tiempo con tratamiento antirretroviral (años)	Tiempo con IP (años)	Tiempo con tenofovir (años)	CD4 Absoluto	CD4 Porcentaje	CD8 Absoluto	CD8 Porcentaje	Carga viral	Ingesta calcio	Creatinina
CD8 Porcentaje	0.07570	0.01936	-0.26782	0.02319	-0.06750	0.03279	-0.11615	-0.41253	-0.72828	0.56647	1.00000	0.35614	-0.03147	0.07092
	0.0572	0.6352	<.0001	0.5651	0.0936	0.4158	0.0038	<.0001	<.0001	<.0001		<.0001	0.6615	0.0796
	632	603	632	618	618	618	618	623	632	622	632	624	196	612
Carga Viral	-0.06710	-0.02194	0.02057	-0.11539	-0.19937	-0.13529	-0.15850	-0.19693	-0.26876	0.13708	0.35614	1.00000	-0.11817	-0.00245
	0.0896	0.5886	0.6035	0.0038	<.0001	0.0007	<.0001	<.0001	<.0001	0.0007	<.0001		0.0923	0.9514
	641	610	640	629	630	630	630	619	628	614	624	641	204	621
Ingesta calcio	0.07759	0.07460	0.04478	0.00820	0.05763	0.08762	0.04757	0.05204	-0.05163	0.02636	-0.03147	-0.11817	1.00000	-0.04802
	0.2664	0.2877	0.5218	0.9076	0.4188	0.2185	0.5046	0.4734	0.4690	0.7196	0.6615	0.0923		0.4984
	207	205	207	203	199	199	199	192	199	188	196	204	207	201
Creatinina	0.06569	0.08745	-0.01989	-0.01847	-0.00669	0.07191	-0.06146	-0.01256	-0.11692	0.10041	0.07092	-0.00245	-0.04802	1.00000
	0.1003	0.0325	0.6191	0.6472	0.8685	0.0745	0.1276	0.7574	0.0037	0.0137	0.0796	0.9514	0.4984	
	627	598	627	616	616	616	616	607	616	602	612	621	201	627

Tabla 31. Correlaciones bivariadas entre las variables independientes.

8.7 Covariables Sociodemográficas. Descriptiva de los pacientes con más de una DEXA

Variable		Total
Menopausia	No	95 (90.48%)
	Sí	10 (9.52%)
Tiempo con menopausia (días)	n	10
	Media (DE)	1886.20 (1482.16)
	Mediana	1970.00
	(Min, Max)	(136.00, 4651.0)
	(P ₂₅ , P ₇₅)	(526.00, 2853.0)

Tabla 32. Menopausia de los pacientes con más de una DEXA.

8.8 Covariables Bioquímicas. Descriptiva de los pacientes con más de una DEXA

Variable		Total
CD4 Nadir	n	387
	Media (DE)	228.09 (145.93)
	Mediana	220.00
	(Min, Max)	(2.00, 807.00)
	(P ₂₅ , P ₇₅)	(107.00, 318.00)

Tabla 33. CD4 Nadir de los pacientes con más de una DEXA.

Variable		Total
CD4 Absoluto	n	375
	Media (DE)	564.11 (286.51)
	Mediana	501.00
	(Min, Max)	(16.30, 1938.0)
	(P ₂₅ , P ₇₅)	(373.00, 699.00)
CD4 Porcentaje	n	381
	Media (DE)	26.98 (8.94)
	Mediana	27.00
	(Min, Max)	(4.00, 54.00)
	(P ₂₅ , P ₇₅)	(20.00, 34.00)

		Total
Variable		
CD8 Absoluto	n	372
	Media (DE)	1030.94 (484.84)
	Mediana	977.50
	(Min, Max)	(193.00, 3408.0)
	(P ₂₅ , P ₇₅)	(680.50, 1280.5)
CD8 Porcentaje	n	379
	Media (DE)	48.42 (11.54)
	Mediana	48.00
	(Min, Max)	(21.00, 79.00)
	(P ₂₅ , P ₇₅)	(39.00, 56.00)

Tabla 34. Linfocitos CD4 y CD8 de los pacientes con más de una DEXA en el momento de la última DEXA.

		Total
Variable		
Creatinina	n	376
	Media (DE)	84.68 (15.93)
	Mediana	83.30
	(Min, Max)	(48.00, 149.70)
	(P ₂₅ , P ₇₅)	(73.00, 94.40)

Tabla 35. Creatinina de los pacientes con más de una DEXA.

		Total
Variable		
Filtrado glomerular	<=60	1 (4.55%)
	>60	21 (95.45%)

Tabla 36. Filtrado glomerular de los pacientes con más de una DEXA.

		Total
Variable		
Porcentaje de tiempo con CV indetectable	< 75%	315 (83.11%)
	>= 75%	64 (16.89%)

Tabla 37. Porcentaje de tiempo con CV indetectable de los pacientes con más de una DEXA.

8.9 Comorbilidades. Descriptiva de los pacientes con más de una DEXA

		Total
Variable		
Hepatitis	No	249 (69.94%)
	Sí	107 (30.06%)

Tabla 38. Hepatitis de los pacientes con más de una DEXA.

		Total
Variable		
Patología concomitante	No	363 (93.80%)
	Sí	24 (6.20%)

Tabla 39. Patología concomitante de los pacientes con más de una DEXA.

8.10 Covariables Farmacológicas. Descriptiva de los pacientes con más de una DEXA

		Total
Variable		
Tiempo con IP (años)	n	392
	Media (DE)	3.76 (3.07)
	Mediana	3.07
	(Min, Max)	(0.00, 13.23)
	(P ₂₅ , P ₇₅)	(1.41, 5.77)

Tabla 40. Tiempo con IP de los pacientes con más de una DEXA en el momento de la última DEXA.

		Total
Variable		
Tiempo con tenofovir (años)	n	392
	Media (DE)	2.38 (2.76)
	Mediana	1.54
	(Min, Max)	(0.00, 29.14)
	(P ₂₅ , P ₇₅)	(0.00, 3.68)

Tabla 41. Tiempo con tenofovir de los pacientes con más de una DEXA en el momento de la última DEXA.

		Total
Variable		
Porcentaje de tiempo con tratamiento antirretroviral	< 25%	10 (2.71%)
	[25 - 50)	47 (12.74%)
	[50 - 75)	100 (27.10%)
	>= 75%	212 (57.45%)

Tabla 42. Porcentaje de tiempo con tratamiento antirretroviral de los pacientes con más de una DEXA.

		Total
Variable		
Tratamiento hormonal sustitutivo	No	383 (98.97%)
	Sí	4 (1.03%)

Tabla 43. Tratamiento hormonal sustitutivo de los pacientes con más de una DEXA.

		Total
Variable		
Tratamiento concomitante	No	340 (87.86%)
	Sí	47 (12.14%)

Tabla 44. Tratamiento concomitante (Corticoides, antiepilépticos, quimioterapia y anticoagulantes) de los pacientes con más de una DEXA.

		Total
Variable		
Calcio	No	376 (95.92%)
	Sí	16 (4.08%)

Tabla 45. Complemento de calcio de los pacientes con más de una DEXA.

		Total
Variable		
Alendronato	No	370 (95.61%)
	Sí	17 (4.39%)

Tabla 46. Alendronato de los pacientes con más de una DEXA.

8.11 Modelo lineal mixto

	Estimación	Error estándar	P-valor
Efecto aleatorio	0.6988	0.1674	<0.0001
Efectos fijos			
Término independiente	1.1128	0.6119	0.0736
Tiempo	-0.0062	0.0008	<0.0001
Edad	-0.0621	0.0161	0.0001

Tabla 47. Estimación de los parámetros para el modelo lineal mixto. Primera DEXA año 2000.

		Estimación	Error estándar	P-valor
Efecto aleatorio		0.6026	0.1749	0.0003
Efectos fijos				
Término independiente		-2.8480	0.6910	0.0001
Tiempo		0.0011	0.0010	0.2895
Tiempo con IP		-0.2044	0.0659	0.0022
Patología concomitante (Categoría de referencia: Sí)	No	1.7838	0.7127	0.0132

Tabla 48. Estimación de los parámetros para el modelo lineal mixto. Primera DEXA año 2002.

	Estimación	Error estándar	P-valor
Efecto aleatorio	0.9195	0.2527	0.0001
Efectos fijos			
Término independiente	0.1405	1.0349	0.8925
Tiempo	0.0003	0.0038	0.9419
Edad	-0.0506	0.0234	0.0325

Tabla 49. Estimación de los parámetros para el modelo lineal mixto. Primera DEXA año 2003.

	Estimación	Error estándar	P-valor
Efecto aleatorio	0.5073	0.1724	0.0016
Efectos fijos			
Término independiente	0.2372	0.9514	0.8049
Tiempo	-0.0029	0.0032	0.3657
Creatinina	-0.0284	0.0111	0.0133

Tabla 50. Estimación de los parámetros para el modelo lineal mixto. Primera DEXA año 2004.

		Estimación	Error estándar	P-valor
Efecto aleatorio		0.4447	0.1830	0.0075
Efectos fijos				
Término independiente		-4.7958	0.8297	<0.0001
Tiempo		-0.0020	0.0068	0.7736
Tiempo con tto. arv. (Categoría de referencia $\geq 25\%$)	< 25%	1.3087	0.3513	0.0006
IMC		0.0942	0.0329	0.0067
CD4 nadir		0.0028	0.0011	0.0157

Tabla 51. Estimación de los parámetros para el modelo lineal mixto. Primera DEXA año 2005.

		Estimación	Error estándar	P-valor
Efecto aleatorio		0.2677	0.1760	0.0641
Efectos fijos				
Término independiente		-0.1486	0.7848	0.8528
Tiempo		-0.0005	0.0047	0.9138
Tiempo con IP		-0.1140	0.0523	0.0444
Edad		-0.0470	0.0171	0.0140
Tiempo con tto. arv. (Categoría de referencia $\geq 25\%$)	< 25%	1.1751	0.5266	0.0403

Tabla 52. Estimación de los parámetros para el modelo lineal mixto. Primera DEXA año 2006 ó 2007.

9 Apèndice

9.1 Encuesta de factores de riesgo asociados a la osteopenia y a la osteoporosis

ENQUESTA FFRR ASSOCIATS A OSTEOPÈNIA/OSTEOPOROSI

Inicials pacient: _____ NHC: _____

Data DEXA: _____

Condicions:

1.- DEXA amb criteris osteopènia/osteoporosi

2.- Si existeix ≥ 1 DEXA \rightarrow agafar la més recent o la primera?

Data enquesta FFRR (dades): _____

Ingesta de calci: _____ mg/dia

Alcohol: Sí ☐ No ☐

Tabac: Sí ☐ No ☐

Nº cigarretes/dia: _____

Exercici físic: Sí ☐ No ☐

(20-30 minuts/dia mínim = sí)

Cafè: Sí ☐ No ☐

Tasses/dia: <2 ☐ 2-5 ☐ >5 ☐

Menopausia: Sí ☐ No ☐

Malalties associades:	Malnutrició	Sí <input type="checkbox"/> No <input type="checkbox"/>	Data: _____
	Endocrinològiques	Sí <input type="checkbox"/> No <input type="checkbox"/>	Data: _____
	Hepàtica	Sí <input type="checkbox"/> No <input type="checkbox"/>	Data: _____
	Renal	Sí <input type="checkbox"/> No <input type="checkbox"/>	Data: _____
	Gastrointestinal	Sí <input type="checkbox"/> No <input type="checkbox"/>	Data: _____
	Artritis	Sí <input type="checkbox"/> No <input type="checkbox"/>	Data: _____
	Càncer hematològic	Sí <input type="checkbox"/> No <input type="checkbox"/>	Data: _____

Ttos concomitants:	Corticoides	Sí <input type="checkbox"/> No <input type="checkbox"/>	Data inici: _____	Data fi: _____
	Antiepil·lèptics	Sí <input type="checkbox"/> No <input type="checkbox"/>	Data inici: _____	Data fi: _____
	Quimioteràpia	Sí <input type="checkbox"/> No <input type="checkbox"/>	Data inici: _____	Data fi: _____
	Anticoagulants	Sí <input type="checkbox"/> No <input type="checkbox"/>	Data inici: _____	Data fi: _____
	Hormones	Sí <input type="checkbox"/> No <input type="checkbox"/>	Data inici: _____	Data fi: _____

Tractament per osteoporosi: Quin? _____

Data inici: _____ Data fi: _____